

PREVISÃO DA VARIABILIDADE DOS FATORES QUÍMICOS E BIOLÓGICOS EM
ÁREA DE RESSURGÊNCIA DE ARRAIAL DO CABO ATRAVÉS DE REDES
NEURAIS

Gilberto Carvalho Pereira

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS
PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA
CIVIL.

Aprovada por:

Prof. Nelson Francisco Favilla Ebecken, D. Sc.

Prof. Luiz Landau, D. Sc.

Prof. Ricardo Coutinho, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2002

PEREIRA, GILBERTO CARVALHO

Previsão da Variabilidade dos Fatores
Químicos e Biológicos em Área de Ressurgência
de Arraial do Cabo Através de Redes Neurais

[Rio de Janeiro] 2002

IX, 78 p. 29,7 cm (COPPE/UFRJ, M.Sc.,
Engenharia Civil, 2002)

Tese – Universidade Federal do Rio de
Janeiro, COPPE

1. Redes Neurais Temporais
2. Previsão
3. Produção Primária

I. COPPE/UFRJ II. Título (série)

“Devemos desejar e procurar o último grau de perfeição, não para nos tornarmos superiores aos outros, mas colocarmo-nos sob os pés de todos”.

À minha mulher Andréa e filhos muito queridos

AGRADECIMENTOS

Ao Prof. Ricardo Rhomberg pelo incentivo dado para que eu fizesse este curso.

Ao amigo Antonio Felipe por sua colaboração.

Ao Prof. Luis Landau por ter percebido através de uma entrevista que mesmo vindo de uma área diferente me encaminhou à minha área de concentração e ao meu orientador.

Ao Prof. Nelson Ebecken pela orientação e oportunidade dada para desenvolver este trabalho.

A todos os colegas e funcionários da COPPE pela assistência e agradável convivência.

Ao Dr. Ricardo Coutinho por ter fornecido os dados inéditos para execução deste trabalho e por valiosas discussões neste período.

Ao CNPq pela bolsa de estudos concedida e a UFRJ pela boa formação acadêmica.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M. Sc.)

PREVISÃO DA VARIABILIDADE DOS FATORES QUÍMICOS E BIOLÓGICOS EM
ÁREA DE RESSURGÊNCIA DE ARAIAL DO CABO ATRAVÉS DE REDES
NEURAS

Gilberto Carvalho Pereira

Março/2002

Orientador: Nelson Francisco Favilla Ebecken

Programa: Engenharia Civil

Modelar a qualidade da água, prever as dinâmicas das populações para definir a saúde do ecossistema é uma função de alguns índices (resiliência, diversidade, produção etc.) e produto de interações de muitas variáveis ambientais.

O acesso às propriedades e processos da zona costeira é o maior tema de discussão no gerenciamento dos sistemas aquáticos.

O objetivo deste trabalho é a avaliação da capacidade de predição das redes neurais quanto as mudanças de clorofila-a como uma medida da biomassa do fitoplâncton. Os dados disponíveis referem-se a uma série temporal média de amostras semanais referentes ao intervalo de 1994 à 2001, e foi coletada em área de influência da ressurgência perto da costa do estado do Rio de Janeiro, sudeste do Brasil.

Lidando, em particular, com uma alta não linearidade entre as variáveis, uma abordagem de algoritmos genéticos foi usada para definir a melhor arquitetura neural e o pré-processamento específico é apresentado.

Os resultados mostram a distribuição dos valores obtidos pelos modelos no curto, médio e longo prazo. Conclui-se que as redes neurais podem com sucesso ser usadas na predição dos complexos fatores impactantes, e o comportamento dos sistemas ecológicos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M. Sc.)

FORECASTING BIOLOGICAL AND CHEMICAL FACTORS VARIABILITY AT
ARRAIAL DO CABO UPWELLING FIELD THROUGH NEURAL NETWORKS

Gilberto Carvalho Pereira

Marsh/2002

Advisor: Nelson Francisco Favilla Ebecken

Department: Civil Engineering

Modelling water quality, forecasting population dynamics to define the ecosystem health is a function of some indices (resilience, diversity, production etc.) and yield of many environment variables interactions. The assesment of properties and processes of costal zone is a major issue in aquatic system management. The aim of this paper is to evaluate neural network capability in predicting changes of chlorophyll-a as a phytoplankton biomass. Available data concern to a weekly medium-term time-series ranging from the end of 1994 to 2001, and it was collected at the Arraial do Cabo near shore upwelling, southeast of Rio de Janeiro state (Brazil). Dealing in particular with high non-linearity among variables, a genetic algorithm approach was used to settle the best neural architecture and the specific pre-processing method presented. The results show the spread of values given by the models in short, medium and long-term predictions. Thus, we conclude that neural networks can successfully be used to improve forecasting of complex impact factors and ecological system behavior.

ÍNDICE

FICHA CATALOGRÁFICA	ii
DEDICATÓROA	iv
AGRADECIMENTOS	v
RESUMO	vi
ABSTRACT	vii
ÍNDICE	viii
CAPÍTULO I – INTRODUÇÃO	1
I.1 – Objetivos	1
I.2 – Complexidade dos Ecossistemas	2
I.3 – Interesses	3
I.4 – Revisão Bibliográfica	4
I.5 – Organização do Trabalho	7
CAPÍTULO II – ATIVIDADE FITOPLANCTÔNICA DE ARRAIAL DO CABO	9
II.1 – Caracterização de Área de Estudo	9
II.2 – O Processo de Ressurgência	11
II.3 – A Atividade Fitoplanctônica	12
II.4 – Os Dados	17
II.4.1 – Tratamento dos dados	24
CAPÍTULO III – IMPLEMENTAÇÃO DO MODELO DE REDES NEURAIIS	25
III.1 – Considerações Gerais	25
III.2 – O Modelo Neural	26
III.3 – Treinamento	29
III.4 – Arquitetura Neural	31
III.5 – Conectividade	32
III.6 – Algoritmo Genético	33

CAPÍTULO IV – ANÁLISE DOS RESULTADOS	37
IV.1 – Previsão	37
IV.2 – Especificações dos modelos	38
IV.2.1 – O Modelo Percptron Multicamadas (MLP)	39
IV.2.2 – Redes com Atraso no Tempo (TDNN)	40
IV.2.3 – Redes com Tempo Adaptável Continuamente (CATNN)	41
IV.4 – Precisão dos modelos neurais	41
IV.5 – Avaliação dos resultados	43
CAPÍTULO V – CONCLUSÕES	47
REFERÊNCIAS BIBLIOGRÁFICAS	51
ANEXO I – MATRIZ DE CORRELAÇÃO	61
ANEXO II – GRÁFICOS	62
ANEXO III – PARÂMETROS E ARQUITETURAS NEURAIIS	72

Capítulo I - Introdução

I.1) Objetivos

As ações do homem têm nas últimas décadas alterado as feições do planeta pela utilização e demanda sobre os múltiplos recursos naturais levantando forte preocupação quanto a sustentabilidade do fluxo dos mesmos e a estabilidade dos ecossistemas. A exploração de um ecossistema e a preservação de sua entidade ecológica representam em geral metas contraditórias. Torna-se cada vez mais imperativo encontrar um meio termo entre os interesses econômicos e as exigências ecológicas, aplicando uma política de manejo racional dos recursos naturais sem destruí-los (Carmouze, 1994). É preciso desenvolver estudos que revelem a dinâmica dos elementos biogênicos. Assim, fazem-se necessários estabelecer métodos que permitam interpretar as flutuações das variáveis ambientais, contribuindo para o entendimento do padrão de comportamento e complexidade dos ecossistemas numa escala de tempo, para futuras previsões.

O objetivo deste trabalho é testar e analisar a capacidade da metodologia de previsão da clorofila-a baseada em redes neurais artificiais como estimativa da produção fitoplanctônica em área influenciada pela ressurgência de Arraial do Cabo como parte integrante de sistemas de monitoração para o apoio a decisão no gerenciamento sustentável dos recursos ecológicos do ambiente costeiro, e o estabelecimento de níveis de segurança. Acessar com precisão o corrente estado dos sistemas aquáticos e predizer seu comportamento futuro requer o desenvolvimento de métodos quantitativos através dos quais possam servir de medidas de sustentabilidade do sistema.

Este trabalho se baseia na hipótese de que os modelos de Redes Neurais são capazes de aprender e modelar a atividade planctônica de Arraial do Cabo a partir dos dados e variáveis levantadas.

I.2) Complexidade dos Ecossistemas

O problema geral de toda análise ecológica e processo de decisão ambiental é a enorme complexidade do ecossistema investigado. Para se construir uma projeção que seja capaz de representar as características mais importantes do estado ambiental, o complexo conjunto de elementos do ecossistema e as múltiplas redes de ações, reações e interações tem de ser considerada de forma aplicável. Um método para alcançar estes objetivos são modelos baseados na utilização de indicadores. Estes são variáveis ou índices que representam, integram e caracterizam a informação incorporada nos dados.

Os ecossistemas são uma classe de sistemas considerados como irreduzíveis (Wolfran, 1984), possuindo uma estrutura hierárquica quanto aos níveis de organização biológica com características próprias de adaptação e organização. Múltiplos objetivos de curto e longo prazo são perseguidos pelas espécies existentes. Esses objetivos podem ser considerados geneticamente pela história de seu processo evolutivo, por essa razão difícil de descobri-lo. Não se pode simplesmente projetar experimentos que revelem as relações que detalham a situação de comportamento de um ecossistema e tentar transferir para outro (Jorgensen, 1999).

O crescimento de organismos fitoplantônicos depende da interação de muitos fatores que são funções do tempo. Mecanismos de “feedback” regulam simultaneamente todos os fatores e taxas, e estes integram entre si e são também funções do tempo em diferentes escalas (Straskraba, 1979). Patten (1997), mostra a importância da história do ecossistema e suas componentes enfatizando a necessidade de uma abordagem dinâmica e sustenta a idéia que nunca se observa a mesma situação duas vezes. O ecossistema apresenta, além disso, um alto grau de heterogeneidade. Todas as suas componentes, principalmente as biológicas estão em constante mudança. Cada situação é diferente de qualquer outra, oferecendo diferentes condições para as várias espécies selecionadas utilizar procedimentos de otimização, colaborando para uma enorme adaptabilidade, tentando obter a maior biomassa para se manterem o mais afastado possível de um equilíbrio termodinâmico.

Processos naturais são controlados por muitas fontes simultaneamente ao nível bioquímico, fisiológico, trófico, habitat, competição intra e inter populacional, e seus efeitos são interrelacionados.

I.3) Interesse

Conhecer e acompanhar a dinâmica dos elementos biogênicos e as principais estruturas tróficas, suas relações com os níveis superiores e com o ambiente físico a longo prazo é de fundamental importância. Este estudo aborda questões relativas a produção primária da atividade fitoplanctônica marinha. Ter uma precisa estimativa deste parâmetro é vital para entender a capacidade de suporte do ecossistema. Conhecer os processos naturais para se regular a entrada de nutrientes na zona costeira, requer conhecimento quantitativo da resposta da comunidade fitoplanctônica que constitui o principal impacto nos processos de eutroficação. Esse conhecimento permite a elucidação da resiliência da comunidade e define os níveis críticos de carga do sistema. Pequenas trocas na produção primária podem produzir uma grande mudança nos níveis tróficos superiores. O fitoplâncton é a principal entidade biológica no ciclo do carbono marinho, desempenhando relevante papel nas trocas do clima global através de sua capacidade de tracionar o dióxido de carbono atmosférico para o meio oceânico para a utilização do mesmo no processo de fotossíntese e na produção de gases que contribuem na formação do núcleo de condensação de nuvens e chuva ácida. Como essas trocas afetam o suprimento de alimentos no meio marinho? O fitoplâncton é a base da rede alimentar, e o maior fator que afeta a qualidade das águas estuarinas, e proporciona ainda uma vantajosa substância para a investigação de processos dinâmicos. A produtividade primária é a taxa de troca da biomassa e reflete a atividade fotossintética da coluna d'água. O fitoplâncton é o principal fator de absorção da radiação solar e entrada de energia nos ecossistemas aquáticos. Portanto, desenvolver técnicas precisas para a estimativa e predição da produção primária torna-se cada vez mais necessário.

Um fato interessante é revelado na bibliografia, onde percebe-se a aplicação de modelos de redes neurais artificiais apenas em dados oriundos de zonas de clima temperado.

I.4) Revisão Bibliográfica

Estudos sobre a composição, distribuição, fenologia e produtividade da comunidade fitoplanctônica da ressurgência de Arraial do Cabo, Rio de Janeiro, tem sido foco de vários projetos de pesquisa. Trabalhos experimentais para avaliação do efeito da predação e competição revelando a estrutura populacional podem ser achados na literatura (Valantin 1984, 1987; Gonzalez-Rodriguez, 1992). A produtividade primária desta área é muito importante para a produção global desta região tropical. A análise da variabilidade da biomassa e sua relação com as condições hidrológicas da oceanografia local contribuem para o melhor entendimento deste ecossistema. Modelos de previsão e produção são ferramentas em potencial para o gerenciamento de corpos d'água sujeitos a processos de eutroficação. As primeiras tentativas de modelagem da produção primária desta área aparecem na literatura com Valantin & Coutinho, 1990 numa abordagem de simular um pico de clorofila durante um evento de ressurgência usando medidas de vento, correntes, temperatura da água, nitrato e clorofila durante cinco dias numa estação fixa com medições a cada quatro horas. O diagrama conceitual inclui como variáveis de estado os valores de clorofila, nitrato e temperatura. Os fluxos usados foram a taxa de crescimento (V_{max}) como uma função do nitrato (Michaelis-Menten), a taxa de mortalidade (grazing) e uma função senoidal do tempo. O modelo faz simulações para três dias e seu melhor resultado considera um fator de advecção. Valantin, 1992 faz um modelo de simulação da distribuição vertical das variáveis, nitrato, biomassa primária e zooplâncton herbívoro para dez dias, sob diferentes condições limitantes de nitrato na camada eufótica. O modelo é aberto a "n" blocos verticais superpostos. A clorofila, zooplâncton e o nitrato formam as variáveis de estado. Luz e pressão do herbivorismo são as funções forçantes no fitoplâncton. Uma taxa de turbulência permite o progressivo transporte do nitrogênio entre os blocos.

Em 1999, Carbonel e Valantin desenvolveram um modelo de acoplamento físico-biológico. O modelo físico descreve as taxas de “momentum” (quantidade de movimento), massa e calor na camada dinâmica, e é forçado pela ação do vento na superfície. O modelo biológico descreve as trocas de suas três componentes (nitrato, fitoplâncton, zooplâncton) e é excitado pela concentração da biomassa primária injetada na camada superficial como consequência dos ventos favoráveis da ressurgência e pela luz produzindo a fotossíntese do fitoplâncton. O acoplamento é estabelecido pela velocidade horizontal, espessura da camada de mistura e a velocidade da ressurgência para as soluções físicas. As equações são resolvidas numericamente no espaço e no tempo por métodos finitos. Entender e quantificar os efeitos do fenômeno da ressurgência é de grande importância para o gerenciamento dos recursos marinhos desta região. Medir diretamente este processo é extremamente complexo e custoso.

Segundo Jorgensen 1999, o desenvolvimento de modelos ecológicos requer o conhecimento do funcionamento do ecossistema, e isso é extremamente importante para se achar o complexo balanço considerando-se os dados disponíveis, o ecossistema e o problema em foco. Para este autor a modelagem ecológica faceia sérios problemas que se deve sobrepujar:

- A insuficiência de dados para desenvolver modelos que permitam prognósticos mais confiáveis;
- A estimativa de parâmetros é freqüentemente o ponto fraco na modelagem;
- Os modelos convencionais não refletem as propriedades reais do ecossistema, particularmente sua adaptabilidade e habilidade para encontrar trocas nas funções forçantes com trocas na composição das espécies.

Diversas abordagens têm sido feitas quanto à solução destes problemas.

Modelos fuzzy têm sido usados para sobrepujar o problema da pobreza da base de dados, o uso da teoria do caos e fractal na modelagem para melhorar a estimativa de parâmetros, o uso da teoria de catástrofe na modelagem como uma tentativa em modelos de troca estrutural (ver Jorgensen,1997), banco de dados de parâmetros ecológicos, o

desenvolvimento de modelos com estrutura dinâmica usando funções objetivo (J, J, J, J, J, & J) levando em conta as propriedades do ecossistema, e o uso de inteligência artificial.

A produção primária normalmente depende não linearmente de fatores externos (luz, temperatura, etc.). A predição de séries temporais produzidas nesses sistemas não pode ser feita com modelos lineares, são manejáveis apenas com métodos não lineares.

Lek (1996), estudando as relações entre a densidade de uma espécie de truta em sítios de desova e as características do habitat compara um modelo de regressão múltipla com outro de redes neurais, onde o último mostra claramente uma melhor performance por sua habilidade de lidar com variáveis não lineares.

Scardi (1996), cita que muitos modelos empíricos tem sido desenvolvidos afim de obter uma estimativa da produção fitoplanctônica. Esses modelos são baseados em regressões da produção do fitoplâncton contra a biomassa e outras variáveis. Scardi compara dois modelos empíricos, um linear mais simples, de formulação geral para a estimativa da produção do fitoplâncton proposto por Cole & Cloern (1984), outro composto com formulação mais elaborada proposto por Hardin (1986) e um modelo de redes neurais, concluindo que os modelos lineares sempre superestimavam os valores de predição quando os valores observados eram baixos e subestimavam quando eles eram altos, enquanto que a rede neural fornecia resultados muito mais precisos. Outra comparação interessante é dada por Barciela, Garcia e Fernandez (1999), onde desenvolvem um modelo convencional de simulação dinâmica e outro de rede neural usados para descrever e prever o principal padrão da variabilidade temporal da produção primária em uma baía afetada por ressurgência (Ria de Arousa, Espanha). O modelo de simulação dinâmica é fisicamente forçado pela radiação solar, temperatura, profundidade da camada de mistura e um índice de ressurgência, desenvolvido usando programação orientada a objetos. O modelo é definido por seis compartimentos biológicos: nanofitoplâncton, microfitoplâncton, microzooplâncton, mesozooplâncton, bactéria e mexilhões cultivados. O modelo de rede neural usou o algoritmo “backpropagation” numa arquitetura de uma ou duas camadas escondidas, usando funções sigmóides e senóides. O modelo de simulação dinâmica contribui para um melhor entendimento do ciclo de matéria através da rede alimentar planctônica, apesar de reproduzir o principal padrão de variabilidade em larga escala, sua

predição potencial é baixa por causa da grande incerteza associada com a estimativa de parâmetros. O modelo neural não fornece informação sobre o funcionamento do ecossistema, mas tem demonstrado ser uma poderosa ferramenta de predições precisas.

Provavelmente a maioria das abordagens de modelagem é baseada sobre uma estrutura de compartimentos, cada compartimento representando um nível trófico ou um grupo taxonômico e as interações expressas por diferentes fluxos ocorrendo entre eles (Odum, 1971).

Resultados de uma análise de sensibilidade de duas redes neurais desenvolvidas para a predição de produção primária em escala regional de um estuário (Chesapeake Bay) nos Estados Unidos é mostrado por Scardi & Hardin (1999).

A aplicação de um modelo neural em escala global para produção fitoplanctônica usando dados de sensoriamento remoto pode ser visto no trabalho de Scardi (2000) no qual a autora sugere novas abordagens para melhorar a performance do modelo. Por exemplo, a utilização de co-preditores permitem uma melhora quando muitas variáveis são selecionadas. Inferir conhecimento teórico do processo biológico por meio de um procedimento de um treinamento forçado (constrained training). Para melhorar a predição do modelo pode-se juntar informações modeladas e reais por meio de um metamodelo, este procedimento ajuda a melhorar a generalização da rede.

I.5) Organização do trabalho

No capítulo II é apresentada uma breve descrição da área de estudo, o mecanismo básico de funcionamento do processo de ressurgência e uma revisão sobre a atividade fitoplanctônica na área estudada. Encontra-se ainda neste capítulo, a caracterização dos dados e o tratamento a eles dispensados.

No capítulo III, encontra-se de maneira resumida uma apresentação dos conceitos teóricos básicos necessários para a implementação do modelo de redes neurais artificiais.

O capítulo IV demonstra o resultado da análise abordando questões relativas à precisão do modelo, prazo de predição e avaliação da rede neural adotada para a série temporal estudada.

Por fim, no capítulo V apresenta-se as principais conclusões e extensões para trabalhos futuros, baseados nas avaliações dos resultados encontrados e nas considerações feitas.

Capítulo II - A ressurgência e a atividade fitoplanctônica

II.1) Caracterização da Área de Estudo

A região de Cabo Frio está localizada na costa leste do estado do Rio de Janeiro. Costões rochosos de formação gnaisse-granito datados do período pré-cambriano dominam a paisagem. Existem poucos rios nesta região, um deles situa-se ao norte do litoral da ilha Rasa onde se encontra uma pequena população de manguezal. O outro se encontra ao sul de Jaconé e corre através do canal de Ponta Negra. A região continental adjacente tem um complexo lagunar bem desenvolvido e apresenta formação de dunas com vegetação característica de restinga (Santos, 1996). O clima é considerado uma transição entre tropical quente e subtropical e é dito como sendo seco se comparado á umidade geral da costa sudeste brasileira. As condições do tempo são influenciadas pela alternância de dois anticlones marítimos, tropical e polar (Barbieri, 1975). As marés são regulares semidiurnas, com amplitudes de 0.8 e 1.0 metros, mas frentes frias podem elevar o nível médio do mar (Yoneshigue, 1985). Duas principais épocas podem ser identificadas:

- a) Primavera-verão com anticiclone marítimo tropical com ventos prevalecendo E-NE;
- b) Outono-inverno com frentes frias polares com ventos SW (Valantin, 1987).

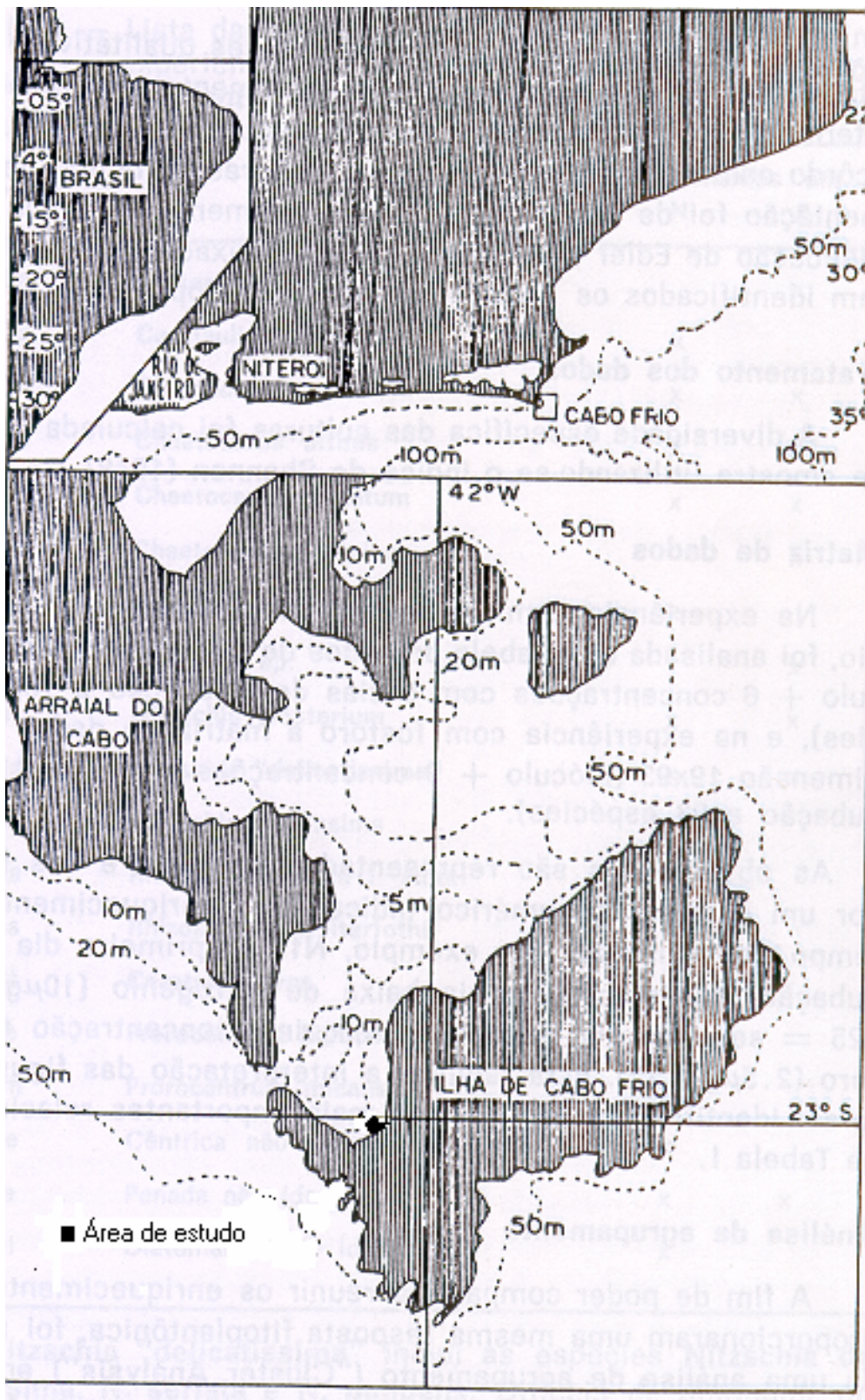


Figura II.1 – Localização da área de estudo

II.2) O processo de ressurgência de Arraial do Cabo

As primeiras evidências da ocorrência da ressurgência aparecem com os trabalhos de Moreira da Silva (1968, 1971 e 1973), além de Valantin (1974). Este fenômeno se manifesta mais freqüentemente entre os meses de outubro e abril, aproximadamente primavera-verão a sudoeste da Ilha de Cabo Frio em Arraial do Cabo. Sua zona de influência pode se estender por mais de cinquenta quilômetros de acordo com sua intensidade (Guimarães & Coutinho, 1996). Até a latitude 22°S o litoral brasileiro apresenta uma larga plataforma continental. Por causa de uma troca na orientação costeira a topografia submarina desta região mostra a isóbata de 50-100 metros se aproximando da costa e na latitude 23°S esta distância é de menos de três milhas (Ilha de Cabo Frio). Nesta mesma latitude a hidrologia própria da região é diretamente dependente das condições meteorológicas, o regime de ventos determina a distribuição das massas d'água. Os ventos E-NE dominantes de um anticiclone marítimo tropical, provocam em razão da força de Coriolis, (transporte de Ekman), uma deriva da água superficial da corrente do Brasil (CB), uma extensão da corrente equatorial sul, que corre em direção sul. A divergência desta corrente da costa contribui para a origem de uma água costeira (AC) derivada da mistura com a corrente de água central do atlântico sul (ACAS) em ascensão de 200-300 metros de profundidade. Este padrão se inverte quando da passagem das frentes frias que proporcionam ventos SW indicadores do fenômeno inverso chamado "subsidência" (retorno para a costa das águas superficiais da corrente do Brasil).

O perfil vertical da temperatura mostra a formação de uma termoclina na zona de encontro das duas massas d'água (superficial e profunda) que flutua de acordo com a intensidade e duração dos ventos (Valantin, 1987).

Paiva (1993), achou uma forte correlação entre o nível médio do mar, o regime de ventos e a temperatura da superfície do mar. Períodos em que o nível médio do mar está baixo estão associados a baixas temperaturas superficiais e ventos predominantes E-NE, sugerindo que o transporte de Ekman é um importante mecanismo que afeta a ressurgência local.

Em resumo, podem ser identificados três principais fatores que aparentemente influenciam a ressurgência.

- a) A topografia;
- b) A posição da corrente do Brasil;
- c) O regime de ventos, associado com o transporte de Ekman.

II.3) A atividade fitoplanctônica

A comunidade do fitoplâncton marinho representa um grupo de diversos organismos que apresentam uma grande atividade metabólica. Essa diversidade que faceia uma aparente estabilidade é enigmática e tem sido estendida além de considerações morfológicas à variações fisiológicas. O plâncton se desenvolve nas massas d'água e, sua distribuição obedece aos movimentos das mesmas dentro de três componentes que caracterizam o meio pelágico: as componentes zonal, meridional e vertical. As duas primeiras definem a distribuição geográfica, ou seja, correspondem respectivamente as variações do plâncton ao longo da costa e ao longo do eixo costa-oceano. A componente vertical é relevante nas áreas onde existe forte turbulência e correntes ascendentes tais como a ressurgência e as áreas de divergência. Uma quarta componente nesta estrutura é a componente temporal, primordial no estudo da dinâmica do plâncton e deve ser bem estabelecida em função do tipo de organismo e do grau de variabilidade dos fatores abióticos a serem modelados (Valantin, 1988). Esta variabilidade provoca diferentes condições oceanográficas afetando de maneira importante todo o ecossistema quanto a sua composição populacional e a produtividade da biota ali presente (Gonzalez-Rodriguez, 1992).

A ressurgência imprime um estresse ao ecossistema que resulta no seu rejuvenescimento com populações abundantes de baixa diversidade específica (Margalef, 1978; Thiriot, 1978).

Durante os meses da estação primavera-verão, a ascensão da Água Central do Atlântico Sul (ACAS) melhor caracterizada por temperaturas inferiores a 18°C, alta

concentração de nutrientes (10 μM de nitrato) e salinidade variando entre 35 e 36 S‰ (Guimarães & Coutinho, 1986), é responsável pela ocorrência e desenvolvimento esporádico de formas planctônicas de origem subtropical, temperada e até batipelágica, enquanto que nos meses da estação outono-inverno referentes ao período de subsidência se encontra o plâncton tropical de águas costeiras de superfície com temperaturas acima de 21°C, salinidade entre 35 e 36 S‰ e baixo teor de nutrientes (nitrato e fosfato abaixo de 1 μM). Este fenômeno constitui para o sistema planctônico um fator geral de variação:

Pela alternância dos movimentos ascendentes e descendentes;

Pela amplitude da componente horizontal;

Pela variação de sua intensidade que caracteriza os ciclos temporais de curtos e longos períodos.

Na variação sazonal as maiores biomassas são registradas durante o verão, os teores de clorofila são superiores a $1\text{mg}\cdot\text{m}^{-3}$, alcançando valores máximos entre 5 e 6 $\text{mg}\cdot\text{m}^{-3}$ com populações microplanctônicas de ordem de 500 a 2 milhões de células por litro (Valantin, 1988). Os ciclos curtos estão ligados a sucessão das fases de ressurgência e subsidência. Durante esse período o fitoplâncton é dominado por populações de *Asterionella glacialis*, *Skeletonema costatum* e *Nitzschia closterium*. Em março-abril, final do período de ressurgência, os máximos são devidos aos gêneros *Chaetoceros* e *Pleurosigma*. O gênero *Rhizosolenia* é predominante em todos os picos invernais sendo acompanhado por uma grande diversidade de espécies de diatomácias entre outras (Valantin, 1988).

A ressurgência é potencialmente importante para a produtividade da biota local (Moreira, 1971; Costa e Haimovici, 1990; Costa e Fernandes, 1993; Coutinho et al., 1989). Os níveis de radiação luminosa em regiões tropicais não são exatamente limitantes para o crescimento algal quando comparados à luz nas altas latitudes (Luning, 1990). Quando a luz se torna favorável ao crescimento algal os nutrientes são consumidos e passam então a serem os fatores limitantes das populações fitoplanctônicas. Variella (1994), cita que além da disponibilidade de luz e nutrientes, outro fator potencial que limita o desenvolvimento das algas é o herbivorismo pelo zooplâncton (grazing).

Os resultados obtidos por Valantin & Coutinho em 1990, a partir da simulação de um máximo de clorofila durante um evento de ressurgência reforçam a hipótese de uma atuação preponderante dos fatores físicos (advecção lateral) e do “grazing” como limitantes do crescimento da biomassa primária em Arraial do Cabo.

A atividade fotossintética do fitoplâncton além de variações sazonais (Harrison & Platt, 1980; Falkowky, 1984), é influenciada pela estratificação da coluna d'água (Gallegos et al., 1983), tamanho das células (Malone, 1977, 1980; Harris, 1983; Takahashi & Biefang, 1983), temperatura (Eppley, 1972; Harrison et al., 1982; Smith et al., 1985), turbulência (Lewis et al., 1984), sendo portanto, fundamental considerar a variação dos parâmetros ambientais. A variação da atividade fitoplanctônica é de grande importância ecológica, uma vez que compromete as medidas de produção diária. Segundo Scuria (1974), se as características fotossintéticas da população variam durante o dia, a produção diária não pode ser diretamente derivada de uma única medida.

Rodriguez & Tenenbaum (1995), observaram a variação diurna da capacidade fotossintética (Pm^B) do fitoplâncton para duas estações de coleta com diferentes graus de influência da ressurgência de Cabo Frio. Os valores de Pm^B ($mg\ C. mg\ Chlor. a^{-1}.h^{-1}$) variam de 0,05 a 23,8 para todos os resultados obtidos e a amplitude máxima ao longo de um dia foi de 0,09 a 4,18. Pm^B apresentou uma relação exponencial com a temperatura e com o nitrogênio. A fração nanoplânctônica representou até 99% do fitoplâncton. A variação da capacidade fotossintética mostrou na maior parte dos experimentos valores máximos no período entre as 10h e 14h. Os valores mínimos ocorreram em uma população constituída por 88% de micropâncton e 12% de nanoplâncton. Segundo Mac Caull & Platt (1977), as variações da fotossíntese em populações naturais resultam de uma oscilação intrínseca do fitoplâncton e de flutuações dos fatores ambientais.

Os resultados de Malone (1971), mostram que quando as concentrações de nitrogênio são limitantes, o nanoplâncton tende a apresentar o máximo de fotossíntese de manhã e à tarde, quando as águas são ricas devido a sua baixa capacidade de estocar nutrientes. Já o micropâncton devido a sua maior capacidade de estocagem, apresenta o máximo à tarde independentemente do regime de nutrientes. Segundo Glover (1980), a

relação entre nutriente limitante e a capacidade fotossintética é variável e altamente dependente da espécie.

Gonzalez-Rodriguez (1994), examinou curvas de saturação de luz (P_m^B) do fitoplâncton de quatro estações da região de Cabo Frio e relatou as trocas das condições ambientais, baseada em 263 experimentos feitos semanalmente durante dois anos. Concluiu que nenhuma diferença na produtividade específica entre as estações foi detectada. A análise de correlação e regressão atribuiu a maioria da variação de P_m^B a temperatura, salinidade, nutrientes e a razão feoficina / clorofila a. Não identificou um significativo padrão sazonal devido a enorme variabilidade temporal de P_m^B mas que esta apresenta alguma tendência. Essa forte variabilidade pode ser explicada pela grande sensibilidade deste ecossistema a trocas repentinas na velocidade e direção dos ventos e a alternância das variáveis hidrológicas e biológicas. A força de estratificação pode também influenciar a produtividade já que isto age sobre a luz e o regime de nutrientes (Jones et al., 1984; Legendre & Demers, 1984).

Em 1992, Gonzalez-Rodriguez, Valantin, André e Jacob, fazem uma análise baseada em diferentes fases de ressurgência e subsidência. Estabeleceram um modelo de análise de componentes principais e regressão dos aspectos biológicos de Cabo Frio. Para condições idênticas de ressurgência, ocorre uma resposta biológica de igual intervalo, permitindo o estabelecimento de um padrão regional. Esses padrões são representados pelos valores de clorofila variando entre 0,5 e 0,6 mg/m^{-3} , e produção primária variando entre 2 e 14 $\text{mg C m}^{-3} \text{ h}^{-1}$. Durante o inverno, na ausência de ressurgência, as águas apresentam uma temperatura, nutrientes, biomassa primária e produção nos mesmos níveis daquelas achadas nas fases de pós-ressurgência. Mesmo na ausência de água profunda, o enriquecimento orgânico parece ser importante, quando comparado às águas oligotróficas da corrente do Brasil. A abundância de matéria orgânica nas águas de inverno, principalmente a maioria oriunda do sedimento ressuspenso do fundo da plataforma continental, constitui a principal fonte de energia, mantendo altos níveis da rede alimentar, especialmente a abundância do zooplâncton (110 mg/m^{-3} /dia, media anual, (Valantin e Moreira, 1978), controlando o crescimento das populações do fitoplâncton.

Num estudo sobre a fertilidade e os fatores nutricionais limitantes da biomassa fitoplanctônica de Cabo Frio, Gonzalez-Rodriguez (1991), afirma que a fertilidade das águas varia com as espécies consideradas e o período do ano. As águas mais ricas em nutrientes minerais (Nitrogênio e Fósforo) não são sempre as mais férteis. No entanto, são as águas de ressurgência que, globalmente, são as mais férteis. O nitrogênio e o fósforo são os dois elementos principais que limitam a biomassa algal; as substâncias complexantes são o terceiro fator limitante. Deduz-se que as algas podem absorver diretamente formas químicas orgânicas dos elementos nutritivos e desta forma, outros elementos podem se tornar limitantes antes dos acima citados.

O conhecimento dos teores mineral da água não é suficiente para conhecer a biomassa máxima potencial que ela pode suportar, pois outros elementos não dosados podem igualmente ser absorvidos e ter papel importante na capacidade de produção algal (Maestrini et al., 1984) assim, influenciando o número de espécies capazes de se desenvolver sobre as condições presentes. Esta propriedade é conhecida como “condicionamento biológico” (Provasali, 1963; Johnston, 1964) e foi colocada em evidência por Barber e Ryther (1969) e Barber et al. (1971). Bansen (1977), mostra que as células no momento da densidade máxima são pobres em clorofila-a, mesmo levando-se em consideração as numerosas variações de ordem fisiológica.

Ryther e Guillard (1959), Goldman e Manson (1962), Charpy e Roubald et al., (1982) assinalaram que a presença de ferro associada com substâncias complexantes pode acelerar a assimilação do carbono, conduzindo a uma sub estimativa do papel do nitrogênio e do fósforo. Além disso, o elemento limitante não é sempre aquele que está presente nas águas nas concentrações mais fracas conhecidas (Smayda, 1970), pois as análises levam, freqüentemente, em conta apenas a parte mineral, enquanto as algas podem utilizar compostos orgânicos (Rayward, 1965; Solorzano e Strickland, 1968 e Anita et al., 1975) e que, de outra forma, o efeito crítico de um elemento depende mais das condições físico-químicas que controlam sua disponibilidade do que sua concentração total (Provasoli, 1963). Este processo é importante para os metais, principalmente o ferro como mostram Huntsman e Sunda, (1980). No entanto, um valor elevado de fertilidade a um dado momento não corresponde forçosamente a um aporte de elementos biogênicos novos,

formas químicas regeneradas participam da produção da biomassa. Os elementos sensu-stricto que limitam a biomassa de Cabo Frio são o nitrogênio, o fósforo, e em menor grau, o ferro e o molibdênio.

II.4) Os Dados

Para se entender a variabilidade dos elementos biogênicos e a resposta futura do sistema, o ideal é examinar estes parâmetros no período de tempo apropriado o que requer uma importante série de dados que refletem a história da área em questão. Neste estudo foram reunidos dados de pouco mais de sete anos, compreendendo o período de novembro de 1994 até dezembro de 2001, fornecidos pelo Instituto de Estudos Almirante Paulo Moreira (IEAPM) e apresentados na tabela I. Vale à pena dizer, que estes dados fazem parte de um projeto de monitoramento ambiental da Ilha de Cabo Frio, mantido por esta instituição, e não foram produzidos no intuito de serem usados em um modelo de rede neural.

Tabela I - Dados Disponíveis

Dados	Unidade	Intervalo
Clorofila	mg/m ³	Semanal
Temperatura	°C	Semanal
Salinidade	‰	Semanal
Oxigênio	ml/l	Semanal
Fosfato	µg/l	Semanal
Nitrato	µg/l	Semanal
Nitrito	µg/l	Semanal
Amônia	µg/l	Semanal

Estes dados consistem de amostragens semanais de água. As coletas foram realizadas na superfície e no fundo da área de estudo e os parâmetros hidrológicos foram

determinados como segue: Para a temperatura foi utilizado um batitermógrafo e um termômetro de inversão acoplado à garrafa de Nansen. Na salinidade foi usado o salinômetro de indução Beckman modelo RS-7C. Com os nutrientes nitrato, nitrito, amônia e fosfato além dos valores de oxigênio usaram-se os métodos descritos em Strickland & Parsons (1972). A clorofila foi determinada através do método espectrofotométrico de Richard & Tompson (1952) e a equação de SCOR-UNESCO (1966).

Os parâmetros físicos, químicos e biológicos das águas coletadas foram anotados em planilhas para posterior tratamento e suas variações são apresentadas na tabela II.

Tabela II – Intervalo (Min./Max.), média e desvio padrão das variáveis físico, químicas e biológicas estudadas.

Temperatura		Salinidade		Oxigênio		Fosfato	
Superfície	Fundo	Superfície	Fundo	Superfície	Fundo	Superfície	Fundo
15.88-26.90	13.00-26.86	32.00-39.14	32.01-39.28	1.97-8.79	0.00-9.00	0.00-1.13	0.00-1.31
22.48	22.14	35.44	35.46	5.32	5.30	0.24	0.28
1.94	2.19	0.67	0.64	0.54	0.64	0.14	0.18

Nitrito		Nitrato		Amonia		Clorofila-a	
Superfície	Fundo	Superfície	Fundo	Superfície	Fundo	Superfície	Fundo
0.00-0.70	0.00-1.35	0.00-8.51	0.00-11.70	0.11-6.61	0.16-11.25	0.00-6.36	0.00-7.09
0.07	0.09	0.53	0.78	1.33	1.44	1.08	1.09
0.09	0.12	0.86	1.57	0.89	1.03	0.94	1.03

Esta base de dados não contém todas as variáveis que podem ser importantes no controle do crescimento fitoplanctônico. Devido à grande transparência das águas estudadas e sua pequena profundidade não se teve a preocupação de incluir informações sobre a incidência de radiação luminosa por entendermos que esta variável neste caso não constitui fator limitante para a produção primária como anteriormente citado por Lunning (1990).

Para ilustrar a compreensão e visualização das relações entre as variáveis envolvidas no crescimento e controle das populações algais o seguinte diagrama conceitual foi elaborado como mostra a figura II.2.

A radiação solar disponível é a principal fonte de energia para a fotossíntese fitoplanctônica e o crescimento destas populações é função da quantidade de radiação incidente e as propriedades ópticas da coluna d'água. A turbidez afeta a penetração da luz e está envolvida na regulação da quantidade de luz do fundo. Da mesma forma a estabilidade da coluna d'água é uma função da turbulência induzida pelo fluxo de correntes e pelo vento. Os dados também não incluem fatores como as interações de competição interespecíficas, vento e fluxo de correntes.

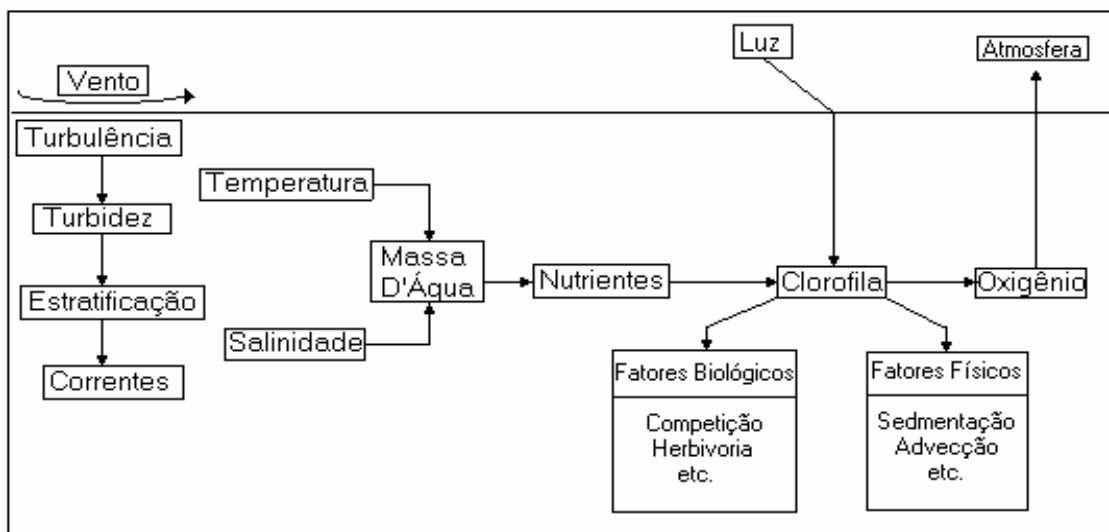


Figura II.2 – Diagrama Conceitual

As massas d'água são primeiramente caracterizadas pela temperatura e salinidade próprias (figura II.3) e transportam os nutrientes que são uma segunda forma de entrada de energia para o sistema fotossintético com posterior produção de oxigênio. Fatores físicos como a sedimentação e advecção; biológicos como a mortalidade, predação, etc. são importantes e controlam simultaneamente o sistema fitoplanctônico.

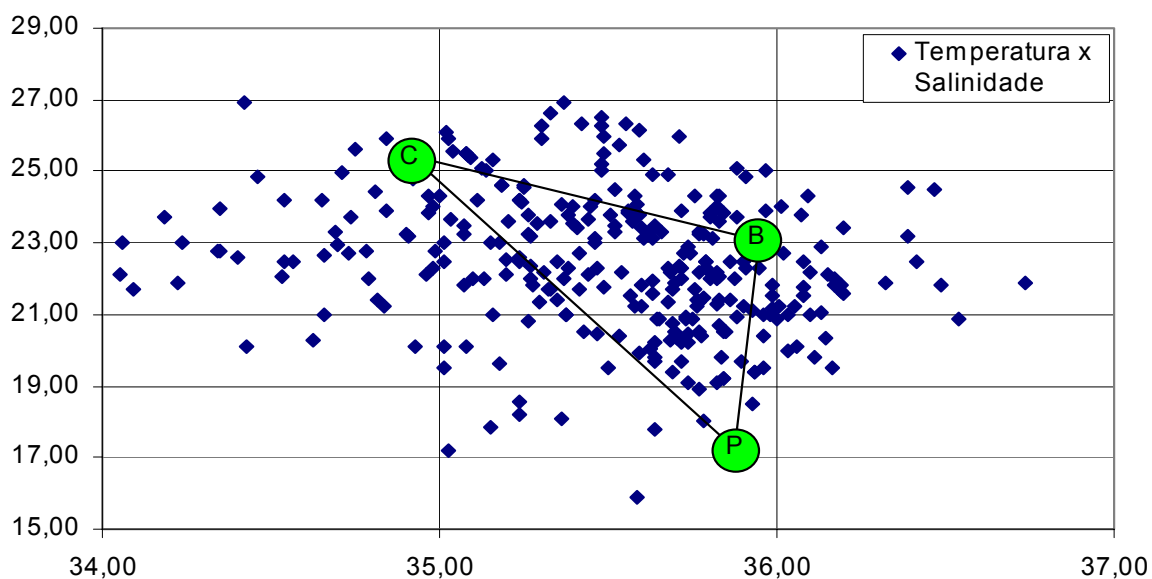


Figura II.3 - Diagrama Temperatura x Salinidade caracterizando os tipos de água: C, costeira; B, corrente do Brasil; P, profunda.

As flutuações das concentrações de clorofila-a (superficial e fundo) ao longo do período de tempo estudado são apresentadas nas figuras II.4a e II.4b e seus valores ultrapassam os verificados por Valantin (1990). Existe uma significativa variação no tempo e uma relativa magnitude de incidência de “blooms” algais que geralmente acontecem nos meses de primavera-verão demonstrando um padrão sazonal dos ciclos de curto período. Verifica-se ainda que o maior valor de clorofila-a ocorreu no fundo, onde tem uma forte incidência em fevereiro de 1996, outra em março de 2000, não sendo detectadas a nível superficial. O ano de 1995 (com grande influencia do fenomeno "el ninõ"), é marcado por picos de pequenas magnitudes não seguindo o padrão do período 1996/1999. Situação semelhante encontra-se no ano 2000. A explicação para tal comportamento pode estar ligada aos ciclos de médio período.

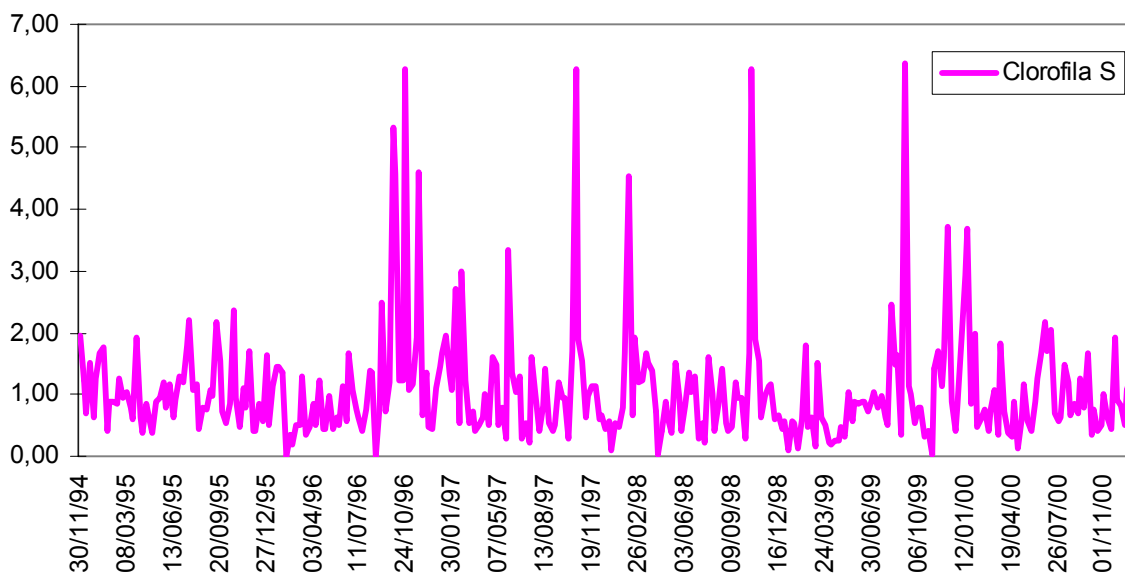


Figura II.4a – Flutuações interanuais das concentrações de clorofila-a na superfície em $\mu\text{g}/\text{m}^3$.

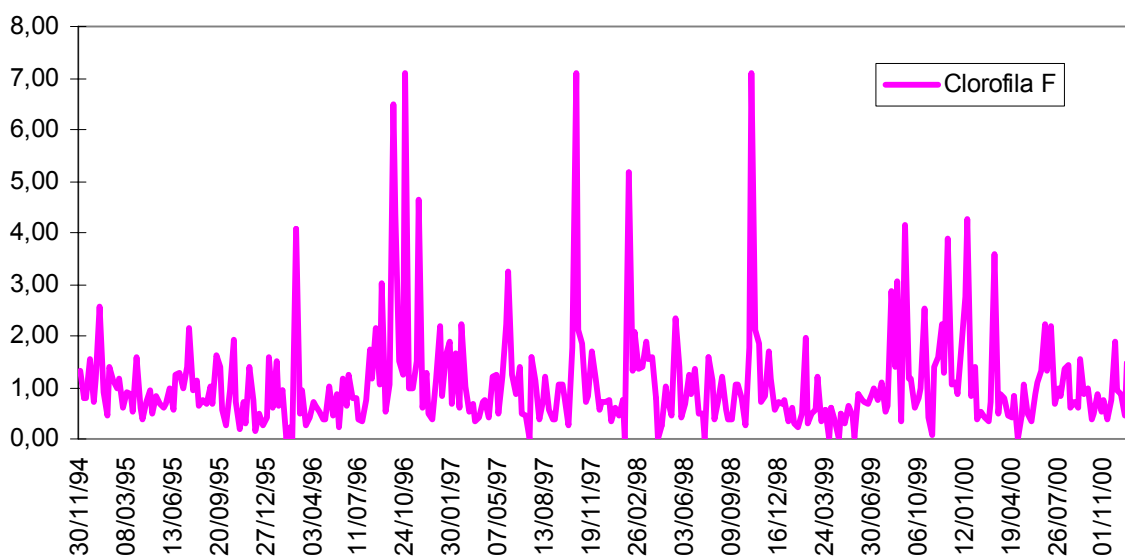


Figura II.4b – Flutuações interanuais das concentrações de clorofila-a no fundo em $\mu\text{g}/\text{m}^3$.

As variações de clorofila relacionadas com as outras variáveis são apresentadas nas figuras II no anexo II. A figura II.4c mostra o perfil dos valores superficiais de clorofila e temperatura. A curva de temperatura tem um comportamento senoidal. Observa-se que os picos de produção primária estão sempre relacionados às temperaturas mais baixas. A figura II.4d (clorofila x temperatura de fundo) novamente mostra um pico de clorofila em

março de 1996 e março de 2000 devido as condições favoráveis da razão nitrogênio/fósforo, o que não ocorreu na superfície. Nas figuras II.4e e II.4f (clorofila x salinidade, superfície e fundo respectivamente) observa-se que a produção primária parece não ser influenciada diretamente pela salinidade, porém se pode notar que no período de agosto/setembro de 1995 são verificados seus mais baixos valores, enquanto que em abril de 1997 temos a situação contrária. Excetuando-se o ano de 1995 que é realmente atípico pode-se notar pequenas perturbações bi-anuais no perfil da salinidade em 1996, 1998 e 2000.

A comparação das figuras II.4g (clorofila e oxigênio de superfície) e II.4h (clorofila e oxigênio de fundo), apresenta uma enorme amplitude e variabilidade para o oxigênio, seu valor máximo é de 9.00 mg/l^{-1} nas águas de fundo e 8.79 mg/l^{-1} na superfície, mas também mostra valores razoavelmente baixos como 1.97 mg/l^{-1} em novembro de 1998 na superfície. Apesar do oxigênio ser um produto da fotossíntese, este não necessariamente acompanha o mesmo padrão da clorofila devido o consumo pelos processos de respiração de todo o ecossistema, processos de oxidação da matéria orgânica e geoquímica, além de sua maior parte ser exportada para a atmosfera.

As figuras II.4i (clorofila e fósforo de superfície) e II.4j (clorofila e fósforo de fundo) quando comparadas mostram a incidência de fosfato nas águas estudadas, evidenciando concentrações mais elevadas no fundo. A ocorrência de picos de clorofila em intervalos de 0.1 à $0.4 \mu\text{M l}^{-1}$ de PO_4 indica que existe uma adequada disponibilidade para o crescimento fitoplanctônico. Sendo o fosfato considerado o segundo nutriente limitante do crescimento algal neste ecossistema, Gonzalez-Rodriguez (1991), a existência de picos deste elemento, não correspondida pela clorofila, pode ser explicado pela razão nitrogênio/fósforo, que quando apresenta valores baixos demonstra a limitação pelo fósforo, caso contrario, valores altos desta razão determina a limitação pelo nitrogênio (ver tabela III).

Das três formas de nitrogênio inorgânico medidas durante o período estudado, foi o nitrato que apresentou o maior valor na superfície e no fundo, $11.70 \mu\text{M l}^{-1}$ e $8.67 \mu\text{M l}^{-1}$ respectivamente. A amônia obedece ao mesmo padrão com valores de $11.25 \mu\text{M l}^{-1}$ na

superfície e $6.1 \mu\text{M l}^{-1}$ no fundo. O nitrito tem a menor participação com concentrações de $0.70 \mu\text{M l}^{-1}$ na superfície e $1.35 \mu\text{M l}^{-1}$ no fundo. A análise dos gráficos referentes às diversas formas de nitrogênio (figuras II.4l, II.4m, II.4n, II.4o, II.4p e II.4q no anexo II) mostra que a ocorrência da clorofila se dá em baixas concentrações de cada uma destas formas químicas, sugerindo que é o seu somatório que contribui para o adequado balanço da razão nitrogênio/fósforo. Também por este motivo pode-se explicar a presença de vários picos de NO_2 , NO_3 e NH_4 não correspondidos pela clorofila como mostrado na tabela III. Outra maneira para se explicar tal comportamento, é que durante o crescimento do fitoplâncton o nitrogênio é consumido muito rapidamente.

Tabela III - Valores de nutrientes, razão N/P, Clorofila e local de ocorrência: S, superfície; F, fundo; S/F, superfície e fundo.

Data	PO4	NO2	NO3	NH4	N/P	Cloro-a	Local
29/3/95	0.25	0.02	0.31	5.97	24.8	1.94	S
12/06/95	1.13	0.11	0.21	1.69	1.78	0.39	S
24/10/96	0.25	0.13	0.29	1.75	8.68	7.09	S/F
29/2/96	0.2	0.04	0.28	1.78	10.5	4.06	F
09/01/97	0.84	1.35	8.67	11.25	25.22	0.83	F
29/4/98	0.1	0	1.59	2	36.92	0.63	S
29/4/98	1.17	0	1.06	1.9	2.54	0.74	F
21/10/98	0.18	0.01	0.65	1.57	12.22	6.28	S/F
09/06/99	0.2	0.34	0	5.92	31.85	0.89	S
15/9/99	0.13	0.04	0.69	1.7	18.81	1.13	S

Os valores da razão N/P muito alto indicam para uma limitação por parte do nitrogênio enquanto valores muito baixos a limitação é dada pelo fósforo. As maiores biomassas ocorrem com valores medianos no período primavera-verão. Um fato interessante acontece em 29/4/98 (início de outono) onde na superfície a limitação é dada pelo nitrogênio, no entanto no fundo se verifica a limitação pelo fósforo. Em 29/2/96 o pico de biomassa se faz presente apenas no fundo já em 15/9/99 um pequeno pico é mostrado só na superfície.

Quando se analisam as figuras II.5a (Temperatura x fósforo superficiais) e II.5c (temperatura x nitrato superficiais) constata-se que ambos nutrientes estão sempre associados a temperaturas mais baixas. O mesmo acontece com a figura II.5e (temperatura x oxigênio superficiais), evidenciando uma correlação negativa, como pode ser visto na matriz de correlações apresentada no anexo I.

II.4.1) Tratamento dos Dados

Uma vez que os dados selecionados foram colocados juntos formando uma única planilha o primeiro procedimento tomado foi o de se verificar a necessidade de padronizá-los em suas respectivas unidades para tornar possível qualquer operação matemática. O segundo passo foi a fase de se conhecer o comportamento da estatística básica, quanto as médias, desvio padrão, valores mínimos e máximos, dados errados, dados faltantes, correlação entre as variáveis etc. A variável salinidade mostrou dois pequenos intervalos de dados faltantes entre 12/9/96 à 1/10/96 e 15/1/98 à 23/4/98 tanto para superfície como para fundo. Através da matriz de autocorrelação (anexo I) pode ser verificado que a variável temperatura tem a maior correlação (negativa) com a salinidade, a qual foi eleita para ser usada no cálculo de regressão linear para o preenchimento dos dados faltantes. A fórmula adotada é a encontrada em Pyle (1999). Na variável oxigênio superficial foram encontrados três valores errados apresentando sinais negativos. Uma consulta ao IEAPM sanou o problema prontamente. A variável nitrito de fundo apresentou apenas dois valores faltantes, neste caso o procedimento adotado foi o de uma média móvel dos últimos quatro valores.

O passo seguinte foi normalizar os dados, segundo a seguinte expressão:

$$Valor\ norm = \frac{Valor\ atual - MIN.}{MAX. - MIN.} \quad (1)$$

Capítulo III - Implementação do modelo de Redes Neurais

III.1) Considerações gerais:

O objetivo deste capítulo é apresentar de maneira resumida os conceitos teóricos básicos para a implementação da técnica utilizada neste trabalho.

Com o aumento do poder computacional adquirido nos últimos anos, o uso de redes neurais artificiais está se tornando cada vez mais comum no desenvolvimento de modelos de predição para sistemas complexos.

As redes neurais são sistemas computacionais que usam a mesma filosofia de modelagem usada pelos modelos baseados em abordagens estatísticas, ou seja, usam um conjunto de dados históricos para aproximar as relações entre as entradas e a variável de saída correspondente. No âmbito da computação neural estão os conceitos de distribuição, adaptação e computação não linear. A estas redes, podem ser atribuídas propriedades que as tornam adequadas para modelar sistemas ambientais como segue:

- Não é necessário o conhecimento prévio da distribuição dos dados de entrada (Burke,1991);
- Os dados podem conter variações sazonais ou cíclicas (Maier and Dandy,1996);
- São altamente não lineares e, funcionam bem quando o conjunto de dados está incompleto ou apresenta algum grau de ruído (Tang et al. ,1990);

O fato de não ser preciso conhecer “à priori” a distribuição probabilística dos dados, fazem delas particularmente interessantes na modelagem de dados ecológicos, pois estes na maioria dos casos, não se apresentam normalmente distribuídos.

Inspiradas nas neurociências, as redes neurais são ditas como sendo uma classe de modelos conexionistas motivadas por forte inspiração biológica, procurando imitar a estrutura e função do cérebro humano. As redes neurais artificiais desempenham uma computação diferente daquela executada pelos computadores convencionais, onde uma simples unidade de processamento (CPU) trabalha de forma seqüencial. As redes artificiais são construídas por um certo número de unidades de processamento simples chamadas de neurônios dispostos em camadas, que as permite lidar com a informação contida nos dados

de forma paralela. O poder da computação neural vem da massiva interconectividade que divide a carga do processamento entre os neurônios e da natureza adaptativa dos seus parâmetros (pesos) que interligam as diversas unidades (sinapses).

III.2) O Modelo Neural

Basicamente, o sistema nervoso consiste em tecidos de células que recebem estímulos do meio, transforma tais estímulos em impulsos nervosos e os transmite para uma área de recepção e integração. Nesta área os impulsos recebidos são interpretados produzindo por sua vez, uma resposta adequada. Todas essas funções são realizadas por meio de um tipo celular altamente especializado denominado neurônio que formam as camadas de tecido do sistema nervoso. Os neurônios variam bastante de tamanho e forma, porém cada um consiste em um corpo celular e um ou mais prolongamentos. A forma de um neurônio depende da função que ele desempenha.

Os modelos conexionistas baseados em redes neurais artificiais possuem características que tentam imitar certos aspectos da cognição humana, evidenciando que, tais modelos podem capturar estas propriedades fundamentais do processamento da informação que é feito no cérebro.

Um neurônio artificial é uma tentativa de modelar o comportamento de um neurônio biológico.

Observando-se o neurônio biológico da figura III.1 pode-se distinguir três partes básicas:

- Os dendritos (entradas) que são prolongamentos nervosos sensíveis aos diversos estímulos que recebe, transmitindo-os para o interior da célula;
- O corpo celular propriamente dito onde, por meio de complexas reações metabólicas, processa-se as informações que a ele chega através dos dendritos. Se o resultado deste processo exceder ao potencial de ação da célula, diz-se que o neurônio foi excitado ou ativado e produzirá uma resposta (reação);

- O axônio, prolongamento de saída, transporta a resposta para o seu terminal ramificado transmitindo o sinal através da sinapse (conexão) com o dendrito do neurônio subsequente.

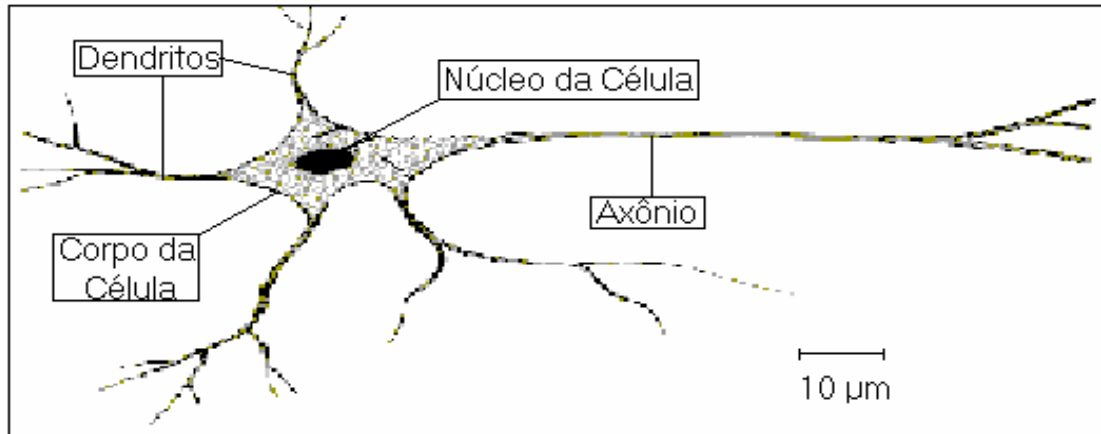


Figura III.1 – Neurônio Biológico

Analogamente, quando se observa o modelo do neurônio artificial na figura III.2, que forma a base do projeto de redes neurais, identificam-se três elementos básicos:

- As sinapses ou conexões, cada uma caracterizada por uma ponderação ou peso (w_k), são representadas por um conjunto de retas que convergem como entradas do referido neurônio;
- Um somador (equivalente ao corpo celular) que soma as entradas (X_1, X_2, \dots, X_n), ponderadas pelos seus respectivos pesos sinápticos. As operações aqui constituem um combinador linear.
- A saída (axônio) representada aqui por uma função de ativação para restringir o sinal dentro de uma certa amplitude.

As funções mais utilizadas são a sigmoideal logística e a tangente hiperbólica cujo valor normalizado da saída do neurônio é escrito no intervalo fechado $[0,1]$ e $[-1,1]$ respectivamente.

Observa-se ainda, a inclusão de uma sinapse chamada “bias” (polarização) representada por X_0 . O bias tem o efeito de aumentar ou diminuir a entrada líquida da

função de ativação dependendo se ele é positivo ou negativo. Sua equivalência com o modelo biológico revela o grau de influência do estado de excitação anterior na saída final do neurônio à chegada de novos sinais.

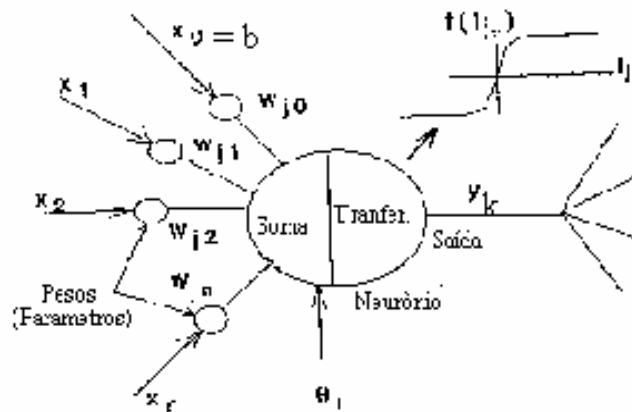


Figura III.2: Modelo do neurônio artificial

O modelo do neurônio artificial funciona então da seguinte forma:

As variáveis de entrada (X_1, X_2, \dots, X_n) que chegam pelas sinapses são multiplicadas pelos seus respectivos pesos (W_1, W_2, \dots, W_n) e então são somados no corpo do neurônio. A este resultado é aplicada uma função de transferência chamada de função de ativação do neurônio, resultando em um sinal de saída Y_k .

Desta maneira, pode-se descrever matematicamente o neurônio U_k segundo as seguintes equações:

$$U_k = \sum W_{kj} X_j \quad (2)$$

$$Y_k = \varphi(U_k + b_k) \quad (3)$$

Onde: X_1, X_2, \dots, X_n são as entradas;

W_1, W_2, \dots, W_n são os pesos sinápticos do neurônio k ;

U_k é a saída do combinador linear;

B_k é o bias;

$\varphi(\bullet)$ é a função de ativação;

Y_k é o sinal de saída do neurônio

Este neurônio é referido na literatura como de Mc Culloch-Pittis desenvolvido em 1943. De acordo com a figura III.3 as redes neurais artificiais são um conjunto de neurônios interconectados dispostos em camadas (camada de entrada), por onde chegam as informações para a rede; um conjunto de neurônios intermediários (camada escondida), onde é feita a maior parte do processamento, e um conjunto de neurônios de saída da rede.

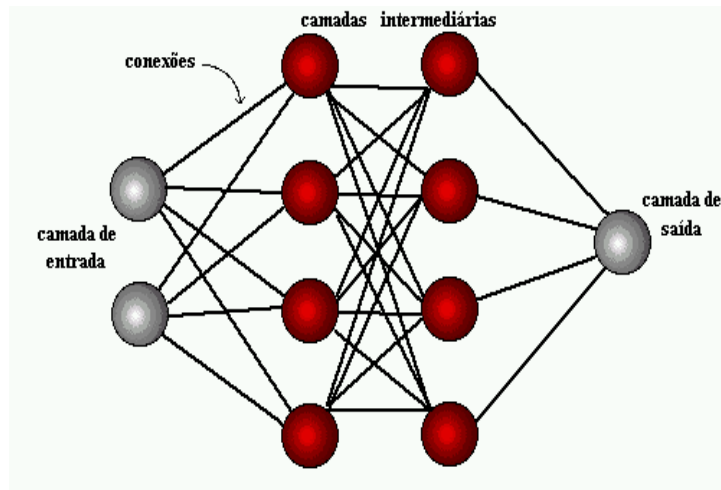


Figura III.3 – Rede Neural Múltiplas Camadas

Existem muitos tipos de modelos de redes neurais encontrados na literatura, para este trabalho será adotada a classificação proposta por N. Sundarajan (1998), baseada no tipo de treinamento, na arquitetura, na conectividade e nos algoritmos de aprendizado.

III.3) Tipos de treinamento

O treinamento de uma rede neural artificial se refere à habilidade de aprender e ao processo de aquisição de conhecimento. Em 1962, Frank Rosenblatt trabalhando com um tipo de rede com neurônios chamados Perceptrons, demonstra que este é capaz de aprender qualquer padrão que ele possa representar.

É importante distinguir entre representação e aprendizagem. Representação se refere a habilidade de simular uma função. Aprendizado requer a existência de um

procedimento sistemático para ajuste de parâmetros (pesos) da rede para produzir a referida função (Vasserman, 1989).

Os métodos utilizados para se definir e ajustar os pesos de uma rede são chamados de algoritmos de treinamento:

Existem três tipos de treinamento:

- Supervisionado;
- Não supervisionado;
- Híbrido ou treinamento por reforço.

No treinamento supervisionado, um agente externo (supervisor) indica à rede a resposta desejada para o padrão de entrada. O ajuste dos pesos se dá de acordo com a diferença entre o valor desejado e a saída da rede. Quando o erro atinge um valor satisfatório diz-se que a rede adquiriu conhecimento e considera-se treinada (figura III.4).

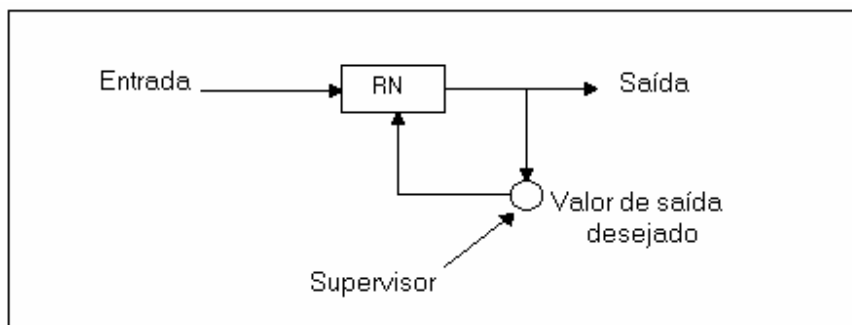


Figura III.4 Treinamento supervisionado

No treinamento não supervisionado, não se dispõe de uma saída alvo, a rede desenvolve a capacidade de representar automaticamente os dados de entrada através de novas classes, interpretando estatisticamente a regularidade dos dados de entrada.

O treinamento híbrido ou treinamento por reforço (figura III.5), pode ser considerado uma variação do treinamento supervisionado. Neste tipo de treinamento não se dispõe de respostas corretas, porém, pode-se saber se as respostas estão corretas ou não. Pode-se dizer que existe um “crítico” que verifica se a resposta da rede é satisfatória, em caso afirmativo as sinapses são reforçadas, caso contrário devem ter menor peso.

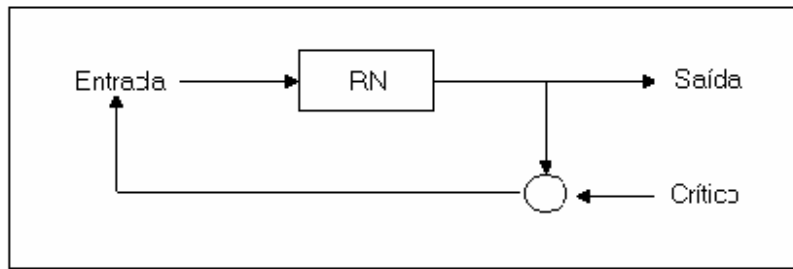


Figura III.5 Treinamento híbrido

III.4) Arquitetura Neural

A definição da arquitetura de uma rede neural artificial é um parâmetro importante na sua concepção, uma vez que ela restringe o tipo de problema que pode ser trabalhado pela rede, seu sucesso ou fracasso.

Uma rede neural pode ser classificada quanto a sua arquitetura como:

- Redes alimentadas adiante ou acíclicas (feedforward). Nestas redes as saídas dos neurônios de uma determinada camada somente se conectam com os neurônios da camada subsequente. A rede é tida como estável e o fluxo de informação é unidirecional, passando sempre da camada de entrada em direção à camada de saída. A figura III.3 anteriormente mostrada, apresenta um exemplo de rede multicamada acíclica.
- Redes recorrentes (feedback).
Nestas redes, pelo menos uma das conexões sinápticas está ligada a algum neurônio de uma camada anterior ou a ele mesmo. A saída global da rede também pode realimentar a própria rede na camada de entrada. São redes melhor adequadas para processamento temporal.

A maneira pela qual os neurônios estão estruturados define a topologia neural e esta intimamente ligada ao algoritmo de aprendizado.

III.5) Conectividade

Apesar dos neurônios biológicos serem mais lentos que os processadores de um computador, eles são melhores numa simples tarefa de reconhecimento de padrão (imagem) por que funcionam de forma paralela e em coletividade usando bilhões de conexões sinápticas.

Por causa das entradas das redes neurais artificiais serem conhecidas, e as funções de ativação serem escolhidas pelo usuário, a saída total pode ser calculada como uma função dos pesos. É através do processo de ajuste desses pesos sinápticos que a rede aprende e adquire conhecimento. Em termos do modelo artificial a conectividade é expressa pelos padrões de ligação entre os neurônios e aos pesos atribuídos a essas ligações.

As possibilidades de ligações entre os neurônios são imensas. A maneira pela qual ocorrem tais conexões e ao número e disposição dos neurônios da rede chama-se topologia.

Uma rede neural pode ser classificada por sua conectividade como:

- Completamente conectada, quando todos os neurônios de uma camada estão conectados a todos os neurônios da camada posterior.
- Parcialmente conectada, caso falte alguma conexão entre os neurônios da camada subsequente.

As conexões podem ser:

- Laterais, entre os neurônios da mesma camada;
- Intercamadas, entre neurônios de diferentes camadas;
- Autoexcitatórias, partem e atingem o mesmo neurônio.

Diferentes tipos de modelos de redes neurais foram testados:

- Perceptron Multicamadas (Multi Layer Perceptron - MLP),
- Redes com Atraso no Tempo (Time Delay – TDNN),

- Redes com Tempo Adaptável Continuamente (Continuous Adaptive Time – CATNN).

Segundo Haykin (2001), quando se utiliza uma rede neural com o objetivo de fazer previsões futuras, o que se deseja é um modelo com capacidade de generalização e, para isto, não se deve treina-las à perfeição.

O processo de definição da arquitetura neural e sua conectividade é inteiramente empírico, especifica a aplicação da mesma e determina o seu sucesso ou fracasso além de consumir algum tempo. Portanto, é necessária a automação desta fase no projeto de redes neurais.

Técnicas de otimização tradicionais podem ser usadas para definir o desenho do sistema neural, mas a maior desvantagem dessas técnicas é que estes algoritmos tendem a terminar em um ótimo local (Jayalakshimi, 2000).

Os algoritmos genéticos são ditos capazes de achar uma otimização global, portanto serão usados neste trabalho para a tarefa da escolha da melhor estratégia de redes neurais.

III.6) Algoritmos Genéticos

Os Algoritmos Genéticos são métodos de otimização e busca desenvolvidos por John Holland em 1975 e, foram inspirados nos mecanismos do processo evolutivo dos seres vivos declarado por Charles Darwing em 1859. De acordo com Darwing, “Quanto melhor um indivíduo se adaptar ao seu meio ambiente, maior será sua chance de sobreviver e gerar descendentes”.

Um Algoritmo Genético (AG) é uma abstração computacional da evolução biológica que pode ser usado para solucionar alguns problemas de otimização (Holland, 1975).

Como técnica de busca e otimização os AGs apresentam:

- Um espaço de busca, onde estão as possíveis soluções para o problema,
- Uma função objetivo, que associa a cada cromossomo uma aptidão (nota).

O espaço de busca é o conjunto de todas as configurações que um cromossomo pode assumir. Os cromossomos por sua vez, são estruturas de dados, geralmente vetores ou cadeias de bits, (uma possível solução para o problema de otimização) que representam os parâmetros da função objetivo.

Cada elemento do vetor cromossomo é chamado de gene, estes determinam as características do ser; a forma mais comum de representa-los é através de codificação binária (0 e 1).

A função objetivo, também chamada de função de adequabilidade (fitness function) quantifica a adaptabilidade de cada cromossomo como uma solução, e é usada como base para seleciona-los para reprodução.

Portanto, o conjunto de cromossomos forma o que se chama de população e as interações do algoritmo genético são chamadas gerações, onde, durante este processo, os operadores genéticos dos AGs agem no material genético.

Existem quatro operadores genéticos (Cole, 1998):

Inicialização, seleção, recombinação e mutação.

- O operador de inicialização é usado para gerar a população inicial para o AG. Esta população deve conter cromossomos que estejam espalhados no espaço de soluções (busca), suprindo o AG de diversificado material genético. A maneira mais fácil de se conseguir isto, é uma seleção randômica dos cromossomos,
- O operador de seleção escolhe os indivíduos para reprodução baseado em suas funções de adequabilidade. Existem diversos métodos para a seleção dentre os quais os mais conhecidos são a Roda da Roleta que utiliza os operadores de recombinação (crossover) e mutação, ou, seleção por Elitismo.

- O operador de recombinação combina o material genético de um ou mais cromossomos pais produzindo uma nova população de cromossomos filhos.
- O operador de mutação introduz novo material genético na população. Em uma codificação binária, por exemplo, simplesmente pode-se mudar um 0 em 1 ou, um 1 em 0 em qualquer ponto do cromossomo.

Na sua forma mais simples, pode-se descrever o Algoritmo Genético como mostra a figura III.6:

- 1 – Codificação do problema original para gerar uma população inicial,
- 2 – Enquanto um certo critério de parada não for satisfeito, faça-se:
 - a – Calcular a aptidão de cada indivíduo na corrente população
 - b – Selecione os indivíduos de alta aptidão e fazer cópias dos selecionados,
 - c – Aplicar os operadores recombinação e mutação criando nova geração.
- 3 – Fim.

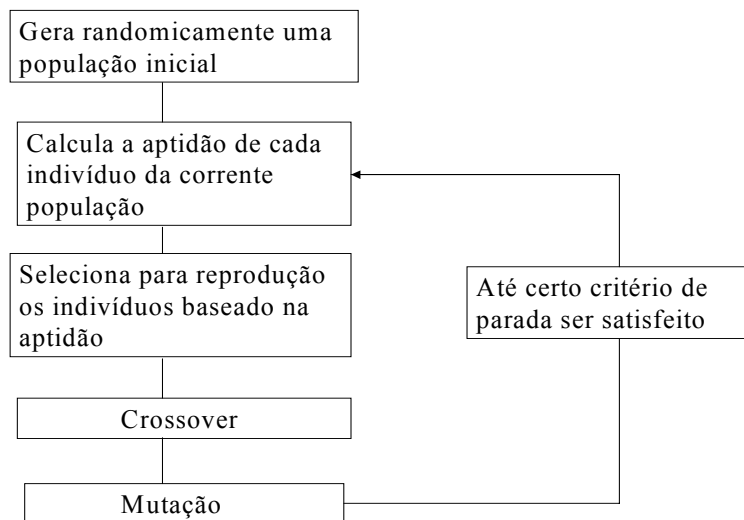


Figura III.6 – Perfil de um simples Algoritmo Genético

Os operadores genéticos de crossover e mutação são os principais mecanismos de busca para explorar regiões desconhecidas no espaço de soluções. Estes combinam a informação genética dos diversos elementos. Como os AGs tendem a reter o material com mais alta adequabilidade e aptidão, a apropriada escolha de adequação aumentará a probabilidade de reter material genético associado com a solução ótima ou próxima desta.

Com o aumento dos recursos computacionais, na década de 90, os Algoritmos Genéticos tornaram-se viáveis na solução de diversos problemas.

Neste trabalho utilizou-se a técnica de Algoritmos Genéticos como processo de otimização do modelo de Redes Neurais para solucionar simultaneamente o seguinte:

- Geração dos pesos de ajuste da rede,
- Geração da arquitetura da rede,
- Análise da rede.

De acordo com os AGs estas informações acham-se representadas em códigos binários e armazenadas nos cromossomos.

Na análise da rede neural, o AG é usado para identificar o melhor padrão de entrada que produz a saída desejada, pesquisa as variáveis mais significativas para obter o melhor resultado.

Cada indivíduo, ou seja, arquitetura de rede neural é treinada e testada quanto ao conjunto de dados (variáveis) disponível obtendo-se uma avaliação da adequabilidade das redes encontradas em função de seu erro, velocidade de convergência etc.

Os parâmetros genéticos e neurais e respectivas variáveis utilizadas por cada arquitetura testada são apresentados no anexo III.

Os detalhes computacionais e aspectos de implementação podem ser encontrados na referência (Biocomp, 1999).

Capítulo IV – Análise de Resultados

Serão apresentados neste capítulo, os modelos de redes neurais utilizados para a predição da atividade fitoplanctônica na superfície e no fundo da área estudada e, os resultados obtidos para os períodos de curto, médio e longo prazo.

Entende-se por previsão de curto prazo o período de quatro meses pertencentes ao intervalo entre 10 de janeiro de 2001 e 18 de abril de 2001. O período de médio prazo constitui um intervalo de oito meses, começando também em 10 de janeiro de 2001 extendendo-se até 18 de julho de 2001. A previsão de longo prazo corresponde à onze meses compreendidos entre 10 de janeiro de 2001 à 14 de novembro de 2001.

IV.1) Previsão

Previsão é uma questão de extrema importância em quase todos os aspectos do mundo atual. Todo o sistema (inclusive o ecossistema) é regido por leis inerentes a sua estrutura e propriedades. Todo o sistema é também composto por subsistemas que possuem diversos estados possíveis.

A forma ideal de se fazer previsões é conhecer as leis que regem o sistema, apurarmos o estado de cada elemento e assim “calcular” o que acontecerá proximamente.

Na prática, no entanto, isto pode ser difícil quando não impossível.

No caso dos ecossistemas, estes possuem uma enorme quantidade de elementos, dos quais uma parte pode não ser conhecida ou acessível. Além disso, propriedades e leis dos ecossistemas podem também ser desconhecidas ou mal compreendidas.

Assim, ao se construir modelos para os ecossistemas, efetua-se apenas uma representação simplificada da realidade.

Como alternativa, usam-se abordagens de previsão baseadas em séries temporais. Nestas séries, espera-se que as leis que governam o sistema, embora não conhecidas ou mal compreendidas, apresentem seu comportamento temporal com algum tipo de regularidade

(padrão). A idéia básica é então, investigar as relações entre as observações que se sucedem no tempo e tentar usar estas relações para extrapolar o comportamento futuro da variável observada.

Na construção de modelos para a previsão de séries temporais é necessário conhecer o relacionamento entre a observação corrente e as anteriores (lag's), para tal, utilizam-se uma função de autocorrelação. Esta função indica como o valor observado está relacionado com seus valores sucessores e anteriores.

IV.2) Especificações dos Modelos

Os modelos de redes neurais artificiais são uma classe de modelos baseados nos dados originais, todavia, uma etapa de pré-processamento mostrou-se crucial como foi demonstrada no capítulo II.

As variáveis selecionadas para cada modelo, seus parâmetros e arquiteturas são apresentados no anexo III.

Para a utilização de redes neurais na predição do caso estudado neste trabalho, a abordagem escolhida foi o treinamento supervisionado seguindo três etapas distintas: A primeira consistiu na apresentação de amostras do conjunto de treinamento do modelo (fase de aprendizagem); para se verificar a aquisição do conhecimento adquirido na fase de aprendizagem, uma segunda etapa, de teste, se faz necessária com outro conjunto de dados (fase de generalização); finalmente na terceira etapa são feitas as predições com um terceiro conjunto de dados (fase de validação). Os conjuntos de dados de treinamento e de teste formam o chamado conjunto total de entradas do modelo. O procedimento aqui utilizado foi escolher aleatoriamente dentro da série temporal de 1994 a 2000, 80% para formar o conjunto de treinamento, e 20% para o conjunto de testes, proporção esta, sugerida por Haykin (2001).

Foi permitido que os modelos evoluíssem até no máximo duas camadas escondidas para evitar o alto esforço computacional.

Curiosamente, a evolução do algoritmo genético utilizado resultou sempre na geração de modelos de redes do tipo MLP.

IV.2.1) O Modelo Perceptron Multicamadas (MLP)

A figura III.3 anteriormente mostrada é um tipo de rede MLP. Tipicamente esta rede consiste de um conjunto de unidades sensoriais que constituem a camada de entrada, um ou mais (dois na figura) camadas ocultas ou escondidas de nós computacionais e uma camada de saída também com função computacional. Uma característica desta rede é a alta conectividade e a utilização do algoritmo de Retropropagação do erro (Error Back Propagation – BP) para os ajustes dos pesos.

Basicamente, a aprendizagem por retropropagação do erro consiste em duas etapas:

Uma etapa de propagação para frente (Forward), e uma etapa de retropropagação, para trás (Backward).

O processo de treinamento se dá da seguinte forma:

Na etapa para frente inicializam-se os pesos com valores aleatórios e pequenos, em seguida, um padrão de entrada do conjunto de treinamento é apresentado aos nós sensoriais da camada de entrada da rede. Este sinal de entrada se propaga para frente através da rede, camada por camada, gerando o cálculo dos valores de ativação ao longo da rede até um vetor saída (\bar{y}) ser produzido. Esta saída é comparada com o vetor de resposta desejada (y) e a diferença entre estes dois vetores obtem-se o vetor erro (e). Durante a etapa de propagação, os pesos sinápticos da rede são todos fixos. O ajuste dos pesos se dá durante a etapa de retropropagação na qual o gradiente do erro quadrático é recursivamente propagado para trás contra a direção das conexões sinápticas segundo a técnica do gradiente descendente. Esta operação de propagação e retropropagação é sucessivamente executada até que se chegue a um erro aceitável ou não se tenha nenhuma melhora.

A função de erro global da saída da rede e o gradiente descendente para a atualização dos pesos podem ser descritos matematicamente segundo as seguintes equações:

$$E = \frac{1}{2} \sum (y - \bar{y})^2 \quad (4)$$

onde:

E – é o erro global, y - é a saída desejada, \bar{y} - é a saída da rede, e

$$\Delta w(t) = -\eta \delta E \delta w + \mu \Delta w(t-1) \quad (5)$$

onde:

η - é a taxa de aprendizagem e, μ - é o valor de momentum.

A correção dos pesos pode ser feita através da apresentação de cada par de entrada – saída do conjunto de treinamento (by-pattern) ou após um certo número de amostras de treinamento ter sido apresentadas à rede chamada de atualização por ciclos ou épocas (by epoch). Segundo Widrow & Sterns (1985), o método de gradiente descendente é o método de atualização de pesos por épocas.

IV.2.2) Redes com Atraso no Tempo (Time Delay – TDNN)

A Rede com Atraso no Tempo (TDNN) foi proposta por Waibel e outros em 1989.

Esta rede tem o funcionamento semelhante à rede MLP, sendo que embute na sua arquitetura memória. Por memória entende-se a introdução aos neurônios da rede de qualquer camada, informações de um período de tempo fixo anteriores aos valores da entrada atual.

As redes com atraso no tempo apresentam as seguintes propriedades:

- Múltiplas camadas com interconexão suficiente entre os neurônios e entre cada camada,
- Capacidade de representar relacionamentos entre os eventos temporais,
- Características reais ou abstrações da aprendizagem da rede não variam com o tempo,
- Os pesos da rede devem ser suficientemente pequenos para extrair regularidades do conjunto de treinamento,

- O processo de aprendizagem não requer um alinhamento temporal preciso dos padrões que serão aprendidos.

Uma rede TDNN simples com um neurônio de entrada e com defasagem de tempo igual a N pode ser entendida como sendo uma rede MLP com $(N+1)$ neurônios de entrada.

O treinamento de rede TDNN é supervisionado, sua arquitetura pode ser “feedforward” ou recorrente, e utiliza o algoritmo de Retropropagação Temporal que apresenta duas características distintas: opera em tempo discreto e apresenta atraso com tempo fixo.

IV.2.3) Redes com Tempo Adaptável Continuamente (Continuous Adaptive Time – CATNN)

Funciona de forma semelhante à rede com atraso no tempo, porém o período de atraso é calculado automaticamente na medida que o treinamento é feito, buscando assim os eventos passados que estão mais correlacionados. Portanto, cada neurônio pode estar associado com períodos de defasagem de tempo diferentes. Esta rede é utilizada principalmente para predição de séries temporais e diagnósticos que evoluam dados de entrada coletados em tempo contínuo. Seu treinamento é supervisionado, sua arquitetura é recorrente e utiliza o algoritmo de Retropropagação Temporal Generalizado, o qual permite que os atrasos no tempo sejam adaptados de uma maneira contínua, do mesmo modo que os pesos são ajustados.

IV.3) Precisão dos Modelos Neurais

Vale a pena destacar, que o treinamento prolongado visando melhorar o desempenho dos modelos de redes neurais artificiais, pode levar a uma superespecialização e, por consequência, a perda da importante propriedade de generalização. O cálculo de precisão do modelo é feito pela diferença entre o valor real e a saída ajustada do modelo. Existem várias maneiras de se calcular a precisão e para isto adotou-se a seguinte notação:

n – número de observações, y – valor real da série no tempo t , \bar{y} - valor predito.

- Erro Médio Quadrático (Mean Square Error – MSE)

Ele é adotado somente quando os erros residuais são independentes é calculado pela seguinte equação:

$$\text{MSE} = \frac{\sum (y - \bar{y})^2}{n} \quad (6)$$

- Raiz do Erro Médio Quadrático (Root Mean Square Error – RMSE)

É uma medida de erro que penaliza mais os erros de maior magnitude.

$$\text{RMSE} = \sqrt{\frac{\sum (y - \bar{y})^2}{n}} \quad (7)$$

- Erro Médio Absoluto (Average Absolute Error – AAE)

Calcula o percentual de erro absoluto em função do número de observações.

$$\text{AAE} = \frac{\sum (y - \bar{y})}{n} \quad (8)$$

- Coeficiente de Determinação (R-Square – R²)

É muito utilizado para medir o ajuste dos dados reais e os definidos pelo modelo. Se o coeficiente é próximo da unidade, diz-se que o ajuste é perfeito, caso contrário, se o coeficiente é próximo de zero o ajuste é fraco.

$$R^2 = \left(\frac{n(\sum y\bar{y}) - \sum y \sum \bar{y}}{\sqrt{n \sum y^2 - (\sum y)^2} \cdot \sqrt{n \sum \bar{y}^2 - (\sum \bar{y})^2}} \right)^2 \quad (9)$$

IV.4) Avaliação dos Resultados

Inicialmente foram geradas seis redes neurais, três modelos de predição para clorofila de superfície e, outros três modelos para a previsão de clorofila de fundo conforme as figuras IV.1 e IV.2 respectivamente.

A observação da figura IV.1 com a previsão da clorofila de superfície mostra que os três modelos foram capazes de boa generalização acompanhando as tendências e a variabilidade da variável estudada.

Os três modelos curiosamente simulam o pico de clorofila no dia 25/7/01 que não aconteceu na realidade, mas conseguem perceber o pico ocorrido em 4/9/01 com relativa perfeição.

Pelos resultados da tabela IV o modelo 5s apresenta o maior coeficiente de determinação (R^2) mostrando um razoável acoplamento das funções desejada e predita. Nota-se ainda que dentre os modelos que simulam a clorofila de superfície é este o modelo de menor RMSE.

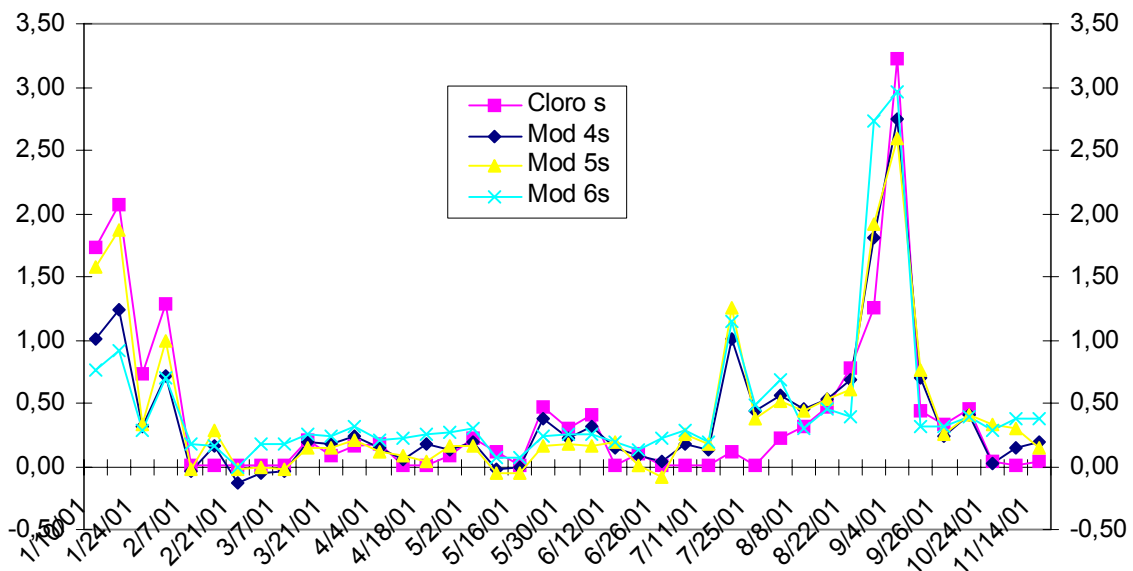


Figura IV.1- Clorofila de Superfície desejada x predita: Modelos 4s, 5s e 6s.

Já quando se examina a figura IV.2 dos modelos de previsão da clorofila do fundo, percebe-se que a performance destes três modelos foi bastante melhor que os modelos de superfície.

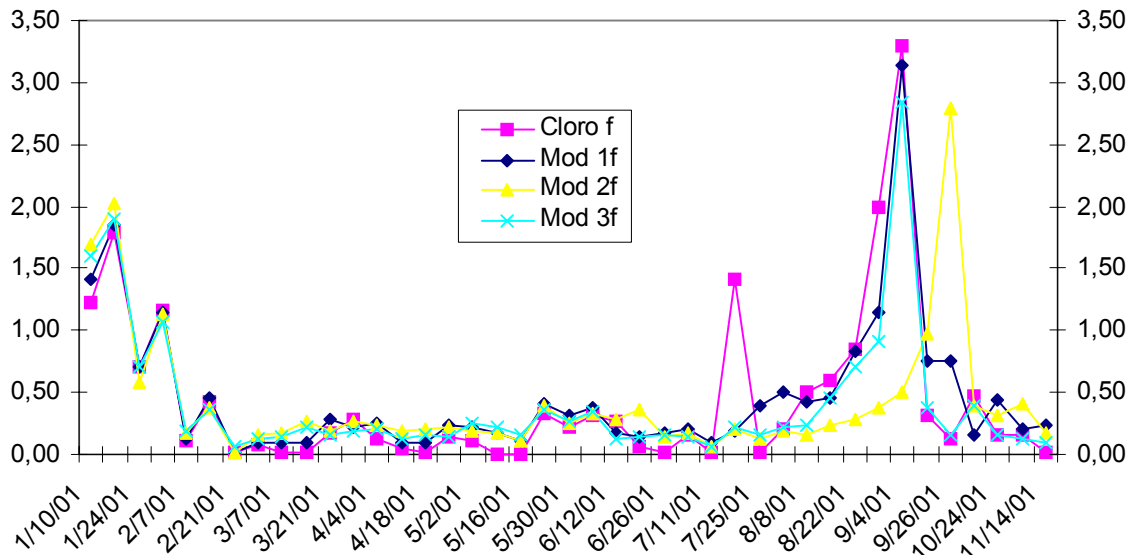


Figura IV.2 Clorofila de Fundo desejada x predita: Modelos 1f, 2f e 3f.

Apesar de nenhum dos modelos de previsão da clorofila de fundo terem sido capazes de prever o pequeno pico ocorrido em 25/7/01 todos apresentam um comportamento muito bom para a predição de curto prazo. No médio prazo pode-se notar algumas diferenças entre os modelos. A simulação dos modelos 1 e 3 no longo prazo conseguem prever o pico de clorofila ocorrido em 4/9/01 porém, o modelo 2 mostra este evento com um certo atraso. Conforme a análise dos erros resultantes dos modelos apresentados na tabela IV, o modelo 1 foi escolhido como o de melhor performance, apresentando a melhor capacidade de generalização, obtendo um alto grau de acoplamento ($R^2 = 0,81222$) e menor erro ($RMSE = 0,08839$), sendo portanto, escolhido para serem feitas previsões de curto e médio prazo.

Tabela IV Estatística dos Erros dos Modelos de Superfície -S, e Fundo -F.

Modelo N°	Previsão	Entradas	MSE	RMSE	AAE	R ²
1-F	4 meses	12	0,00185	0,04303	0,03136	0,95246
1-F	8 meses	12	0,00230	0,04797	0,03466	0,94470
1-F	11 meses	12	0,00781	0,08839	0,05609	0,81222
2-F	11 meses	10	0,00850	0,09218	0,05312	0,79669
3-F	11 meses	11	0,00784	0,08857	0,04912	0,81668
4-S	11 meses	11	0,04404	0,20985	0,12743	0,18889
5-S	11 meses	13	0,01155	0,10745	0,07585	0,71829
6-S	11 meses	14	0,01824	0,13504	0,09015	0,57999

O gráfico da figura IV.3 de previsão de médio prazo do modelo 1 mostra que esta rede é capaz de perceber as variações do sistema fitoplanctônico com bastante precisão. Isto também se vê na figura IV.4 com a previsão de curto prazo onde os acoplamentos tem um valor bastante elevado ($R^2 = 0,95242$ para curto prazo e, $R^2 = 0,94470$ para médio prazo).

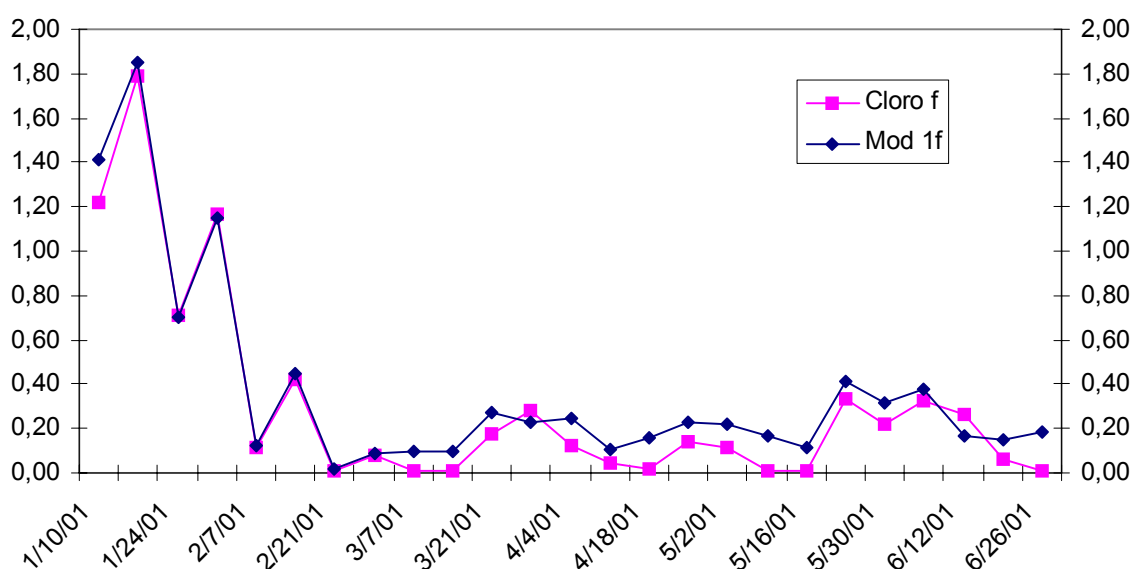


Figura IV.3 Previsão de médio prazo de clorofila de fundo, desejada x predita, Modelo 1

É interessante notar que os modelos não são muito precisos quando se observa valores muito baixos de produção primária, este fato pode estar relacionado com a qualidade dos dados. Na verdade, valores observados de produção primária não devem ser considerados como “valores verdadeiros” por que são freqüentemente afetados por grandes erros de medição (Eppley, 1980). Isto implica que quando estes valores baixos são levados em consideração, o ruído contido no sinal é normalmente alto e conseqüentemente a previsibilidade é muito baixa. As redes neurais não fogem a esta regra.

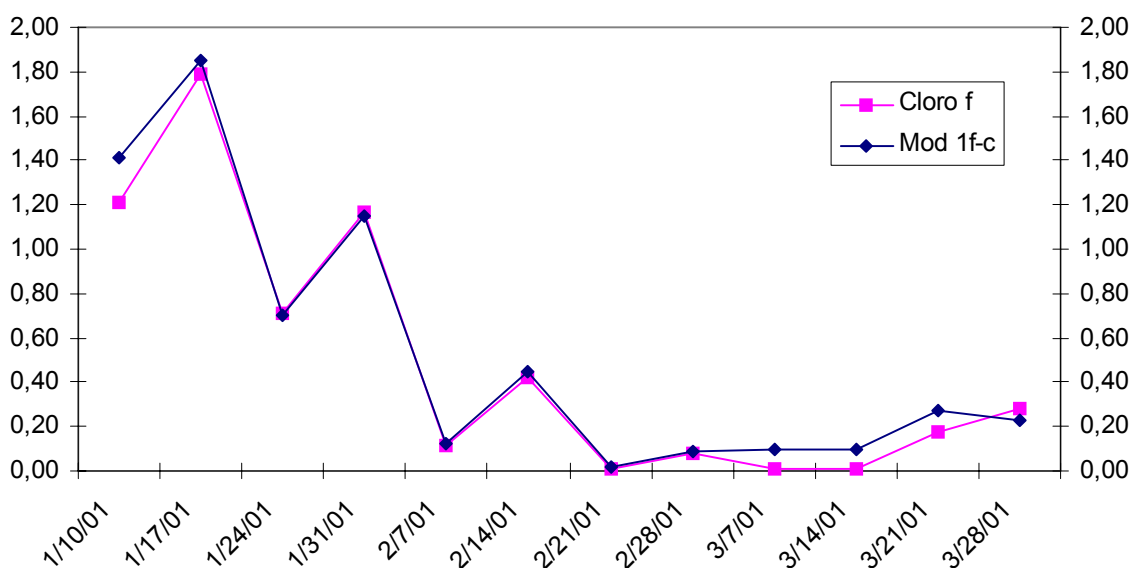


Figura IV.4 Previsão de Curto Prazo da Clorofila de Fundo, desejada x predita , Modelo 1

O fato de a rede ter um melhor desempenho com valores mais altos de produção primária demonstra o importante papel destes valores no direcionamento do treinamento do modelo sendo estes valores muito mais relevantes do que aqueles com baixos valores.

Pela experiência adquirida neste trabalho fica fácil perceber que os modelos gerados diminuem seus erros se forem treinados para previsões de prazos mais curtos

Capítulo V – Conclusões

O problema geral de toda a análise ambiental e todo o processo de decisão é a enorme complexidade dos ecossistemas.

Para se construir uma projeção capaz de representar as características mais importantes do estado ambiental, o complexo conjunto de elementos do ecossistema e as múltiplas redes de ações, reações e interações têm de ser consideradas.

Um método para isto é a utilização de indicadores. Estes são variáveis ou índices que representam, integram e caracterizam a informação incorporada no conjunto de dados. Os indicadores devem fornecer suficiente informação se uma dada condição ambiental está ou não envolvida com os objetivos de uma política de desenvolvimento sustentável.

Mudanças ambientais pelo aumento de carga de nutrientes provocam uma relativa eutroficação e alteração da razão N/P, conseqüentemente, a componente biológica ao nível fitoplantônico responde significativamente com um aumento de biomassa.

Como citado por Margalef (1978) e outros, a eutroficação provoca um aumento da biomassa mas também uma drástica diminuição das espécies existentes, isto ocorre por que certos organismos possuem mecanismos de captura dos nutrientes melhor adaptado a funcionar num nível de concentração mais elevado.

Os sistemas aquáticos podem ser considerados como dominados por plâncton, o que implica, que estas populações devem fornecer valiosa informação no acesso às condições ambientais.

As vantagens de usar o fitoplâncton como indicador incluem.

- O plâncton fornece a mais notável indicação de eutroficação na zona costeira. Mudanças na concentração de nutrientes pode resultar em mudanças da estrutura e função da comunidade. A população dos

produtores primários planctônicos é a que apresenta o menor tempo de resposta as mudanças ambientais.

- Mudanças na produção primária planctônica conseqüentemente afetam os níveis tróficos superiores.
- O uso da clorofila-a é fácil e tem relativo baixo custo para programas de monitoramento.
- O plâncton tem geralmente um curto ciclo de vida, e alta taxa de reprodução fazendo dele um valioso indicador de impactos no curto prazo.

Porém, como todas as populações, existem desvantagens associadas com o uso do plâncton como segue:

- A identificação taxonômica do fitoplâncton pode ser difícil e custosa em termos de tempo.
- A competição com macrófitas, altas taxas de respiração, e o aumento da herbivoria do zooplâncton pode impedir o aumento da biomassa fitoplanctônica resultante do enriquecimento dos nutrientes. Essas razões sugerem que para a investigação o fito e zooplâncton devem ser usados em conjunto como indicadores biológicos.
- O fitoplâncton pode produzir “blooms” e as causas não serem conhecidas.

A modelagem mecânica baseada na representação de ecossistemas como estruturas fixas e parâmetros fixos negligencia um aspecto básico dos sistemas biológicos, sua capacidade de adaptação e auto-organização.

Neste trabalho foi empregada a metodologia de redes neurais artificiais para testar sua capacidade de previsão da variabilidade de resposta biológica frente às mudanças ambientais dos fatores físicos e químicos.

Demonstrou-se que a área estudada é muitas vezes um ambiente de diferentes graus de eutroficação e limitação de nutrientes.

Comprovou-se que as redes neurais podem aprender com sucesso o comportamento da série temporal estudada.

Em termos de modelos de redes neurais, encontrou-se o modelo MLP como sendo o mais adequado neste estudo.

As diferentes arquiteturas, topologia e conjunto de variáveis apresentadas no anexo III demonstram que o mais importante para uma aplicação bem sucedida desta classe de modelo é um ajustado e representativo banco de dados para o problema abordado.

Cabe ressaltar que nem sempre o melhor modelo ajustado representa a melhor previsão (Roitman, 2001).

O modelo de redes neurais são representados por vários parâmetros (peso das conexões), funções de ativação não lineares, e as unidades da camada escondida (que capturam as características) aprendem de forma distribuída.

A precisão das predições dos modelos decresce com o aumento do período de predição.

O modelo 1 mostra uma arquitetura recorrente o que lhe confere comportamento dinâmico.

Para facilitar a aplicação das redes neurais existem diversos softwares que permitem a geração de códigos da rede treinada em qualquer linguagem.

Explicitar o conhecimento adquirido pela rede neural implica no uso dos chamados algoritmos de extração de regras.

Como perspectivas de trabalhos futuros estão:

- Uso de modelos neurais com variáveis que representam as espécies chaves de outros níveis tróficos.

- Estudo e modelagem das interações plâncton – necton – bentos.
- Identificação das espécies planctônicas por modelos neurais no tratamento de micro fotografias (Imagens)
- Identificação por meio de modelos neurais das principais populações fitoplanctônicas dada uma certa condição ambiental.
- Uso de algoritmos de extração de regras.
- Transformação de regras neurais em funções de pertinência fuzzy.
- Implementação de um sistema híbrido neuro-fuzzy para gerenciamento ambiental levando também em consideração variáveis sociais e econômicas.
- Utilização de dados de sensoriamento remoto.
- Previsão de pesca em relação aos níveis de produção primária.
- Previsão dos efeitos do fenômeno "el ninõ", etc.

REFERÊNCIAS BIBLIOGRÁFICAS

ANITA, N. J., BERLAND, B.R., BONIN, D.J. and MAESTRINI, S.Y., 1975, “Comparative evaluation of certain organic and inorganic sources of nitrogen for photothrophic growth of marine algae”. *J. Mar. Biol. Ass. U.K.*, v.55, pp. 519-539.

BANSEN, K., 1977 “Determining the carbon-chlorophyll ratio of natural phytoplankton.” *Mar. Biol.*, v. 41, pp. 199-212.

BARBIERI, E.B., 1975. “Ritmo climático e extração de sal em Cabo Frio”. *Revista Brasileira de Geografia*, v. 37, pp.23-109.

BIOCOMP SYSTEMS, INC., 1999, Neurogenetic Optimizer, Redmond, WA, USA.

BARCIELA, R.M., GARCIA, E.; FERNANDEZ, E., 1999, “Modelling primary production in a coastal embayment affected by upwelling using dynamic ecosystem models and artificial neural networks”. *Ecol. Model.*, v.120, n.2-3, pp. 199-211.

BRAGA, A. P., LUDERMIR, T.B., CARVALHO, A.C.P.L.P., 2000, “LTC” – Livros Técnicos e Científicos Editora S.A.

CARBONEL, C.A.A., VALANTIN, J.L., 1999. “Numerical modelling of phytoplankton bloom in the upwelling ecosystem of Cabo Frio (Brazil)”. *Ecological Modelling*, v.116, pp.135-148.

CARMOUZE, J. P., 1994, *O Metabolismo dos Ecossistemas Aquáticos: Fundamentos teóricos, métodos de estado e análises químicas*. 1 ed. Editora Edgard Blücher: FAPESP.

CHAKRABORTY, K., METHROTRA, K., RANK, S., 1992, “Forecasting the Behavior of Multivariate Time Series Using Neural Network”. *Neural Networks*, v.5, pp. 961 – 970.

CHARPY.R., CHARPY, L.J., MAESTRINI, S.Y., 1982, “Fertilité des eaux côtières Nord-patagoniques: facteurs limitant la production du phytoplankton et potentialités d’exploitation myticole”. *Oceanol. Acta*, v.5, n.2, pp.179-188.

COLE, B.E., CLOERN, J.E., 1984, "Significance of biomass and light availability to phytoplankton productivity in San Francisco Bay". *Mar. Ecol. Prog. Ser.*, v.17, pp.15-24.

COLE, M. R., 1998, *Clustering with Genetic Algorithms*. Tese de D,Sc., University of Western Austrália.

COSTA, P.A.S., Fernandes, F. C., 1993. "Seasonal and spatial changes of cephalopods caught in the Cabo Frio (Brasil) upwelling ecosystem". *Bulletin of Marine Science*, v.52, pp.751-759.

COSTA, P.A.S., HAIMOVICI, P., 1990, "A Pesca de Polvos e lulas no litoral do Rio de Janeiro". *Ciência e Cultura*, v.42, pp. 1124-1130.

DE SÁ, C.B.P., 1996, *Previsão de Séries Temporais Epidemiológicas Utilizando Redes Neurais*. Tese de M.Sc., COPPE/UFRJ, Rio de Janeiro , RJ, Brasil.

DUARTE, C. M., AGUSTÍ; S., NONA, S. R., 2000 "Response of a Mediterranean phytoplankton community to increased nutrient inputs: a mesocosm experiment". *Mar. Ecol. Prog. Ser.*, v.195, pp. 61-70.

EPPLEY, R.W., 1980, *Estimating phytoplankton growth rates in central oligotrophic oceans*. In: Falkowski, P.G., Editor, *Primary Productivity in the Sea*, Plenum Press, New York, pp. 231 – 242.

EPPLEY, R.W.,1972, "Temperature and phytoplankton growth in the sea": *Fish. Bull.* v.,70, pp. 1063-1085.

FALKOWISK, P.O., 1984, "Physiological responses of phytoplankton to natural light regimes".

GALLEGOS, C.L., PLATT, T., HARRISON, W. B., IRWIN, B., 1983. "Photosynthetic parameters of arctic marine phytoplankton: vertical variations and time scale of adaptation". *Limnol. Oceanogr.*, v.28, pp. 698-708.

- GLOVER, H.E., 1980, "Assimilation numbers in cultures of marine phytoplankton". *J. Plank. Res.*, v.2, pp.69-79.
- GOLDMAN, C.R., MANSON, D.T., 1962, "Inorganic precipitation of carbon in productivity experiments utilizing carbon 14". *Science*, v.136, pp.1049-1050.
- GUIMARÃES, M. A., 2000, *The influence of environmental factors on the seasonal dynamics of Ulva spp. and Sargassum sp. In Cabo Frio upwelling region of Brazil*, Ph.D. dissertation, University of Miami, Florida, USA.
- GUIMARÃES, M. A., COUTINHO, R., 1996, "Spatial and Temporal variation of benthic marine algae at the Cabo Frio upwelling region, Rio de Janeiro, Brazil". *Aquatic Botany* v.52, pp. 283-299.
- HARDIN, L. V. Jr., MEESON, B.W., FISHER, T.R. Jr., 1986, "Phytoplankton production in two east coast estuaries: photosynthesis-light functions and patterns of carbon assimilation in Chesapeake and Delaware Bay". *Estuar. Coast Shelf. Sci.*, v. 23, pp. 773-806.
- HARRIS, G. P., 1983, *Mixed layer physic and phytoplankton population studies in equilibrium and non-equilibrium ecology*. In F.E. Round & D.J. Chapman (eds). Progress in Phycological Reserch. 2. Biopress. England: 1-46.
- HARRISON, W.G., PLATT, T., 1980, "Variations in the assimilation number of costal marine phytoplankton: Effects of environmental covariates". *J. Plankton Res.*,v.2, pp. 249-260.
- HARRISON, W.G., PLATT, T., IRWIN, B., 1982, "Primary production and nutrient assimilation by natural phytoplankton population of the eastern Canadian artic". *Can. J. Fish. Aquat. Sci.*, v. 39, pp. 335-345.
- HAYKIN, S., 2001, *Redes Neurais, princípios e prática*, 2^a edição, Rio de Janeiro, Bookman.

HOLLAND, J. 1975, *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

JAYALAKSHIMI, G. A., MUTHARASU, D., RAJARAM, S., PANDIYAN, S. G., KANNIAPPAN, P., 2000, "An Evolutionary Programming Approach to Evolve the Architecture of Artificial Neural Networks". <http://www.ise.nus.edu.sg/proceedings/apros2000/fullpapers/28-06.htm>

JOHNSTON, R., 1964, "Sea water, the natural medium of phytoplankton. II. Trace metals and chelation, and general discussion". *J. Mar. Biol. Ass. U.K.*, v. 44, pp. 87-109.

JONES, J.L., GOWEN, R. L., TETT, P., 1984, "Water column structure and summer phytoplankton distribution in the Sound of Jura Scotland". *J. exp. Mar. Ecol.* v.78, pp.269-289.

JORGENSEN, S. E., 1999, "State-of-the-Art of ecological modelling with emphasis on development of structural dynamic models". *Ecological Modelling*, v. 120, n.2-3, pp. 75-96.

JORGENSEN, S. E., BERNARD, R., 1977, "The application of model with dynamic structure to simulate the effect of mass fish mortality on zooplankton structure in lago de Annone". *Hydrobiologia*, v. 356, pp. 87-96. GEOBASE.

JORGENSEN, S.E., 1994, *Fundamentals of Ecological Modelling*, 2 ed., Elsevier.

LEGENDRE, L., DEMERS, S., 1984, "Towards dynamic Biological oceanography and limnology". *Can. J. Fish. aquat. Sci.*, v. 41, pp.2-19.

LEK, S.; DELACOSTE, M., BARAN, P., DINOPOULUS, I., LAUGA, J., AULAGNIER, S., 1996, "Application of neural networks to modelling nonlinear relationship in ecology". *Ecol. Model.*, v. 90, pp. 39-52.

LEWIS, N.R.; CULLEN, J.J., PLATT, T., 1984, "Relationship between vertical mixing and photoadaptation of phytoplankton: similarity criteria". *Mar. Ecol. Prog. Ser.*, v. 10, pp.1-6.

LI, B. L., CHARNOV, E. L., 2001, "Diversity-Stability relationship revisited: scaling rules for biological communities near equilibrium". *Ecol. Model.*, v. 140, pp. 247-254.

LÜNNING, K., 1990, *Seaweeds their environment, biogeography and ecophysiology*. New York, Wiley-Interscience.

MACCAULL, W.A., PLATT, T., 1977, "Diel variation in the photosynthetic parameters of coastal marine phytoplankton". *Limnol. Oceanogr.* V. 22, pp. 723-731.

MAESTRINI, S.Y., DROOP, M.R., BONIN, D.J., 1984, *Teste of algae as indicator of sea water quality: Prospects*, pp 133-188. In L.E. Shubert (ed.), *Algae as ecological indicators*, Academic Press, London.

MAIER, H. R., DANDY, G. C., 1996, "Neural Network models for forecasting univariate time series". *Neural Netw. World.* v. 5, n. 96, pp. 747-771.

MAIER, H. R.; GRAEME, C. D.; MICHAEL, D. B., 1998, "Use of artificial neural networks for modeling cyanobacteria *Anabena* spp. in the River Murray, South Australia". *Ecol. Model.*, v. 105, pp. 257-272.

MAIER, H. R.; MICHAEL, D. B., 1997, "Modelling cyanobacteria (blue-green algae) in the River Murray using artificial neural networks". *Mathematics and Computers in Simulation.*, v. 43, pp. 377-386.

MALONE, T.C. 1971, "Diurnal rythms in netplankton and nanoplankton assimilation ratios". *Mar. Biol.* V.10, pp. 285-289.

MALONE, T.C., 1977, "Light-Saturated photosynthesis by phytoplankton size fraction in the New York Bight. USA. *Mar. Biol.*, v. 42, pp. 281-294.

MARGALEF, R., 1978, "Lês types biologiques de phytoplancton consideres comme dès alternatives de survie dans um milieu instable", *Ocoanol. Acta.*, v. 1, n. 4, pp. 493-509.

MARTIN, A.P.; KELVIN, J., RICHARDS, FASHAM, M. J. R., 2001, “Phytoplankton production and community structure in an unstable frontal region”, *Journal of Marine Systems*, v. 28, n. 1-2, pp. 65-89.

MOREIRA DA SILVA, P. de C., 1973, “A ressurgência de Cabo Frio”. *Publicações Inst. Pesq. Mar.* v. 78, pp.1-56

MOREIRA DA SILVA, P. de C., 1968, “O fenômeno da ressurgência na costa meridional brasileira”. *Publicações Inst. Pesq. Mar.*, v. 24, pp. 1-38.

MOREIRA DA SILVA, P.C., 1971, *Fertilization on the sea as a by-product of an industrial utilization of deep water*. In: Fertility of the sea. Costlow, J.D. (ed.). Gordon & Breach Sci. Publ. New York, Paris, T I,II,

NADJA, M.L.S., VALANTIN, J.L.; BASTOS, C.T.B., 1988, “O microfitoplanctom do litoral fluminense (estado do Rio de Janeiro): Lista de espécies e aspectos ecológicos”. *Bol. Inst. Oceanogr.*, S. Paulo, v. 36, n. ½, pp. 1-16.

NEUROGENÉTIC OPTMIZER, User’s Guide, Biocomp Systems, Inc, Rednord, WA, 1999.

ODUM, E.P., 1971,: *Fundamentals ol Ecology*. Sanders, Philadelphia, P.A.

PAIVA, A.M., 1993, *Estudo das variações do nível médio do mar em Arraial do Cabo- RJ*. In. Technical Report IEAPM.18-25.

PATTEN, B. C., 1997, “Bear model for Aironduck National Park”. *Ecol. Model.* v. 100, pp.11-42.

PLATT, T., JASSBY, A. D., 1976, “The relationship between photosynthesis and light for natural assemblages of costal marine populations. *J. Phycol.*, v.12, pp. 421-430.

PROVASALI, L., 1963, *Organic regulation of phytoplankton fertility*, In M. M. Hill (ed.) *The Sea*, Vol. 2, Intersc. Publish., New York.

PYLE D., 1999, *Data Preparation for Data Mining*, Morgan Kaufmann Publishers, Inc. San Francisco, California.

RECKNAGEL, F.; MARK, F.; PIA, H.; KEN-ICHI, Y., 1977, “Artificial neural network approach for modeling and prediction of algal blooms”. *Ecol. Model.* v. 96, pp. 11-28.

REFENES, A.N., AZEMA-BARAC, M., CHEN, L., KAROUSSOS, S.A., 1992, “Currency Exchange Rate Prediction and Neural Networks Design Strategies”. *Journal of Neurocomputing and Applications*, v.4, n.2, pp.417 – 427.

RODRIGUEZ, E. G., 1991, “A ressurgência de Cabo Frio (RJ, Brasil): Fertilidade e fatores nutricionais limitantes da biomassa fitoplanctônica”. *Rev. Brasil. Biol.*, v.51, n.2, pp. 471-486.

RODRIGUEZ, E. G., 1994, “Yearly variation in primary productivity of marine phytoplankton from Cabo Frio (RJ, Brazil) region”. *Hydrobiologia*, n. 294, pp. 145-156.

RODRIGUEZ, E. G., TENEBBAUM, D. R., 1995, “Variação diurna da capacidade fotossintética do fitoplânctom de Cabo Frio (RJ, Brasil)” *Acta Biológica Leopoldina*, v. 17, n.1, pp. 35-52.

RODRIGUEZ, E. G., VALANTIN, J.L., ANDRÉ, D.L., JACOB, S.A., 1992, “Upwelling and downwelling at Cabo Frio (Brazil): Comparison of biomass and primary production responses”. *Journal of Plancton Reserch.*, v. 14, n. 2, pp. 289-306.

ROITMAN, V. L., 2001, Um Modelo Computacional de Redes Neurais para Predição do Índice de Desemprego Aberto. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

RUMELHART, D.E.; HILTON, G.E., WILLIAMS, R.J., 1986. “Learning representations by back-propagation error”. *Nature*, v. 323, pp. 533-536.

RYTHER, J.H., GUILLARD, R.R.L., 1959, “Enrichment experiments as a means of studying nutrients limiting to phytoplankton production”. *Deep Sea Res.*, v. 6, pp. 65-69.

SALVADOR, E., 2000, "Costal upwelling in the eastern Gulf of California". *Oceanologia Acta*, v. 23, n. 6, pp. 731-744.

SANTOS, D.P., 1996, *Distribuição espacial e temporal de propágulos de macroalgas marinhas e sua relação com o recrutamento no costão na região de Arraial do Cabo e adjacências*. Tese de D.Sc., Universidade de São Paulo, S.P., Brasil.

SCARD, M., HARDING, L. W. Jr., 1999, "Developing an empirical model of phytoplankton primary production: a neural network case study". *Ecol. Model.* v.120, n. 2-3, pp. 213-223.

SCARDI, M., 1996, "Artificial neural networks as empirical models for estimating phytoplankton production". *Mar. Ecol. Prog. Ser.* v. 139, pp.289-299.

SCARDI, M., 2000, *Neural networks of phytoplankton primary production*. In Lek S/ Guegan J-F [eds.], *Artificial Neural Networks: Application to Ecology and Evolution*, Springer-Verlag, Berlin/Heidelberg.

SMATH, J.C., PLATT, T., LI, W.K.W., HORNE, E. P. W., HARRISON, W. G., SUBBA, RAO, D.W., IRWIN, D. W., 1985, "Arctic marine photoautotrophic picoplankton". *Mar. Ecol. Prog. Ser.*, v. 20, pp. 207-220.

SOLORZANO, L., STICKLAND, J.D.H., 1968, "Polyphosphate in sea waters". *Limnol. Oceanogr.*, v. 13, pp. 515-518.

STRASKRABA, M., 1979, "Natural control mechanisms in models of aquatic ecosystem". *Ecol. Model.*, v. 6, pp. 305-322.

STICKLAND, J.D.H., PARSONS, T. R., 1968, "A practical handbook of sea water". *J. Fish. Res. Bd. Can. Bull.*, v. 167, pp. 311

STUMM, W., MORGAN, J., 1981, *Aquatic Chemistry*, Wiley Interscience, New York.

TAKAHASHI, M., BIEFANG, P. K., 1983, "Size structure of phytoplankton biomass and photosynthesis in subtropical Hawaii". *Mar. Biol.*, v.76, pp. 203-211.

VALANTIN, J.L., MOREIRA, A.P., 1978, “A matéria orgânica de origem zooplanctônica nas águas de ressurgência de Cabo Frio. (Brazil)”. *Nac. Acad. Brás. Ciênc.*, Rio de Janeiro, v. 50, n.1, pp.103-112.

VALANTIN, J.L., 1984, “Analyse des parameters hydrobiologiques dans la remontee de Cabo Frio (Brésil)”. *Marine Biology*, v. 82, pp. 259-276

VALANTIN, J.L., 1988, *A dinamica do planctom na ressurgência de Cabo Frio-RJ*. Inst. Pesqui. Mar., Rio de Janeiro. Coletânea de trabalhos, In: F.P. Brandini (editor) Memórias de III EBP Curitiba, 1989.

VALANTIN, J.L., 1992, “Modelling of the vertical distribution on marine primary biomass in the Cabo Frio upwelling region”. *Ciência e Cultura (Journal of Brazilian Association for the Advancement of Science)*v. 44, n.2-3, pp. 178-183.

VALANTIN, J.L., COUTINHO, R., 1990, “Modelling maximum chlorophyll in the Cabo Frio (Brazil) upwelling: a preliminary approach”. *Ecol. Modelling.*, v. 52, pp. 103-113.

VALANTIN, J.L., 1974, *O plankton na ressurgência de Cabo Frio (Brasil). II Primeiras observações sobre a estrutura física, química e biológica das águas da estação fixa(período 04/02 à 16/04/73)*. Inst. Pesq. Mar. Rio de Janeiro.

VALANTIN, J.L., ANDRÉ, D.L., JACOB, S.A., 1987, “Hydrobiology in Cabo Frio (Brazil) upwelling: two-dimensional structure and variability during a wind cycle”. *Continental Shelf Research*. v. 7, n. 1, pp. 77-88.

VARIELLA, I., 1994, *Marine Ecological Process*, Springer Verlag.

WASSERMAN, P., 1989, *Neural Computing: Theory and Practice*. Van Nostrand Reinhold, New York.

SMAYDA, T.J., 1970, “Growth potential Bioassay of water masses using diatoms culture. Phosphorescent Bay (Puerto Rico) and Caribbean waters”. *Helgoland. Wiss. Meeresunters.*, v. 20, pp. 172-194.

WIDROW, B., STEARNS, S.D., 1985, “ Adaptive Signal Processing, Englewood Cliffs, NJ.: Prentice – Hall.

WOLFRAM, S., 1984, “Cellular automata as models of complexity”. *Nature*, v. 311, pp. 419-424. INSPEC Compendex.

YONESHIGUE, Y., 1985, *Taxonomie et 'cologie des algues marines dans la région de Cabo Frio (Rio de Faneiro, Brésil)*. Ph.D. dissertation, Université de Marseille, France.

Anexo I

	TEM_S	TEMP_F	SAL_S	SAL_F	O2_S	O2_F	F04_S	F04_F	NO2_S	NO2_F	NO3_S	NO3_F	NH4_S	NH4_F	CLOR_S	CLOR_F	N.P.S	V.P.F
TEM_S	1,00	0,90	-0,17	-0,20	-0,06	-0,18	-0,33	-0,28	-3,40	-0,35	-0,25	-0,28	0,04	0,02	-0,26	-0,27	0,15	0,13
TEMP_F	0,90	1,00	-0,17	-0,21	-0,11	-0,12	-0,31	-0,37	-3,39	-0,47	-0,27	-0,50	-0,01	-0,11	-0,23	-0,22	0,12	0,13
SAL_S	-0,17	-0,17	1,00	0,78	0,01	0,39	0,15	0,13	3,10	0,10	0,04	0,05	-0,05	-0,04	0,03	0,06	-0,14	-0,14
SAL_F	-0,20	-0,21	0,78	1,00	0,02	0,13	0,16	0,16	3,10	0,10	0,04	0,06	-0,05	-0,03	0,04	0,06	-0,15	-0,15
O2_S	-0,09	-0,11	0,01	0,02	1,00	0,34	0,08	0,04	3,08	0,05	0,3	0,07	0,00	0,02	-0,02	-0,01	-0,01	0,00
O2_F	-0,18	-0,12	0,09	0,13	0,04	1,30	0,07	0,04	3,05	0,04	0,04	0,00	-0,05	-0,04	0,16	0,15	-0,11	-0,09
P04_S	-0,33	-0,37	0,15	0,18	0,06	0,37	1,00	0,68	3,46	0,39	0,32	0,18	-0,02	0,03	0,07	0,01	-0,47	-0,42
P04_F	-0,28	-0,37	0,13	0,16	0,04	0,34	0,68	1,00	3,36	0,51	0,33	0,54	-0,01	0,09	0,07	0,03	-0,23	-0,35
NO2_S	-0,40	-0,39	0,10	0,10	0,06	0,35	0,46	0,36	1,00	0,75	0,38	0,26	0,08	0,15	0,11	0,08	-0,13	-0,13
NO2_F	-0,35	-0,47	0,10	0,10	0,06	0,34	0,39	0,51	0,75	1,00	0,43	0,56	0,11	0,34	0,08	0,04	-0,05	0,00
NO3_S	-0,25	-0,27	0,04	0,04	0,13	0,34	0,32	0,33	3,38	0,43	1,00	0,59	-0,06	0,09	0,05	0,04	0,07	0,03
NO3_F	-0,28	-0,50	0,05	0,06	0,07	0,30	0,18	0,54	3,26	0,56	0,59	1,00	0,03	0,19	0,06	0,05	0,04	0,13
NH4_S	0,04	-0,01	-0,05	-0,05	0,00	-0,35	-0,02	-0,01	3,08	0,11	-0,06	0,03	1,00	0,81	0,00	-0,01	0,44	0,45
NH4_F	0,02	-0,11	-0,04	-0,04	0,02	-0,34	0,03	0,09	3,15	0,34	0,09	0,19	0,81	1,00	-0,01	-0,04	0,39	0,43
CLOR_S	-0,26	-0,23	0,03	0,04	-0,02	0,16	0,07	0,07	3,11	0,08	0,05	0,06	0,00	-0,01	1,00	0,86	-0,13	-0,13
CLOR_F	-0,27	-0,22	0,06	0,06	-0,01	0,15	0,01	0,03	3,08	0,04	0,04	0,05	-0,01	-0,04	0,86	1,00	-0,07	-0,09
N.P.S	0,16	0,12	-0,14	-0,14	-0,01	-0,11	-0,47	-0,28	-3,10	-0,06	0,07	0,04	0,44	0,39	-0,10	-0,07	1,00	0,74
V.P.F	0,18	0,10	-0,14	-0,14	0,00	-0,39	-0,42	-0,36	-3,10	0,00	0,08	0,13	0,46	0,46	-0,10	-0,08	0,74	1,00

Matriz de correlação

Anexo II

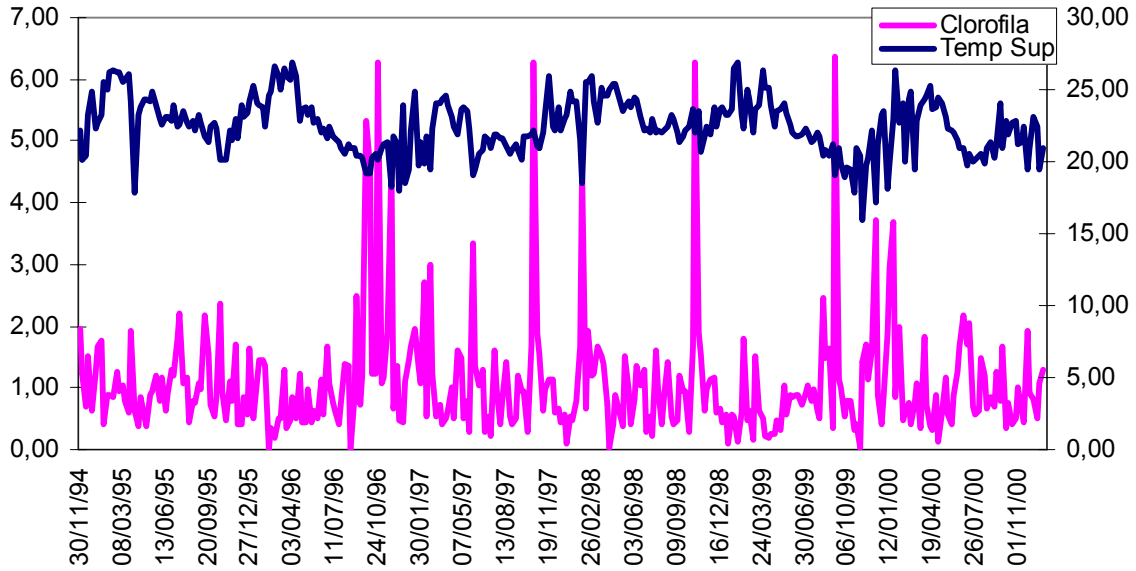


Figura II.4c – Flutuações interanuais da clorofila x temperatura na superfície.

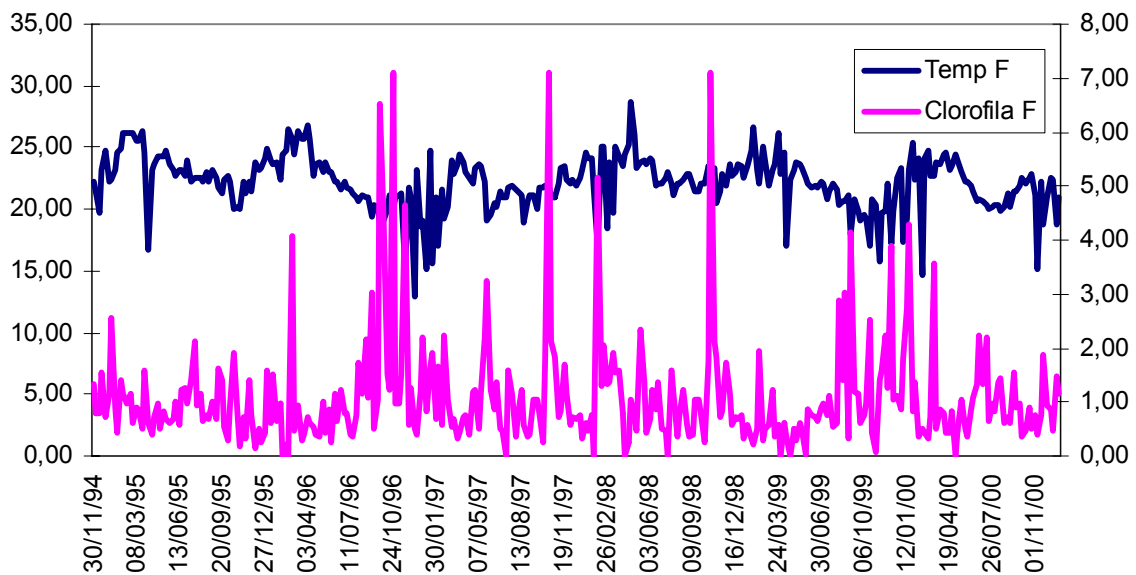


Figura II.4d – Flutuações interanuais da clorofila x temperatura no fundo.

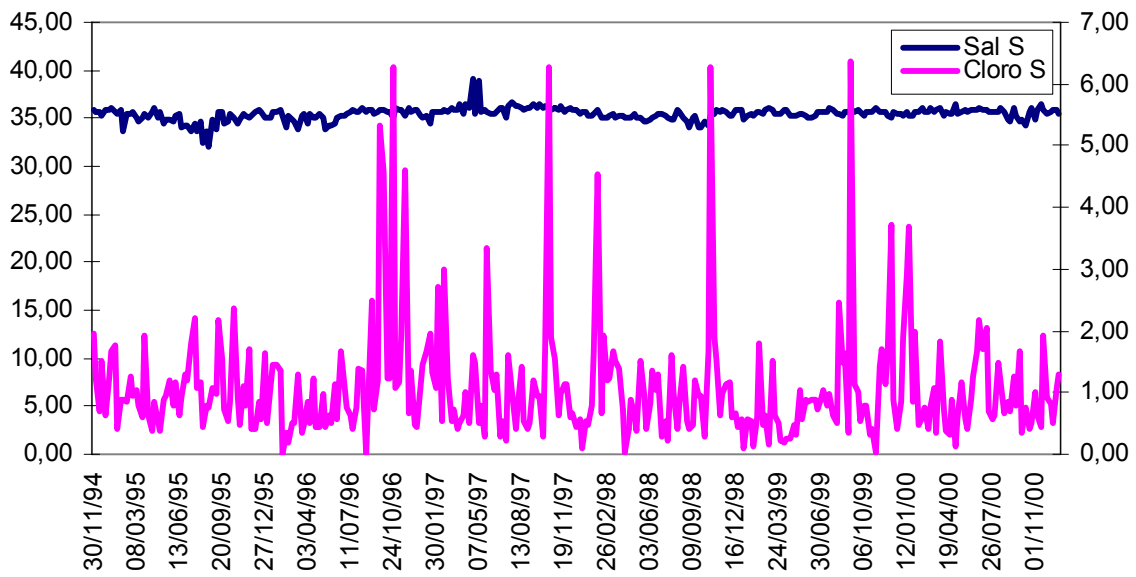


Figura II.4e – Flutuações interanuais da clorofila x salinidade na superfície.

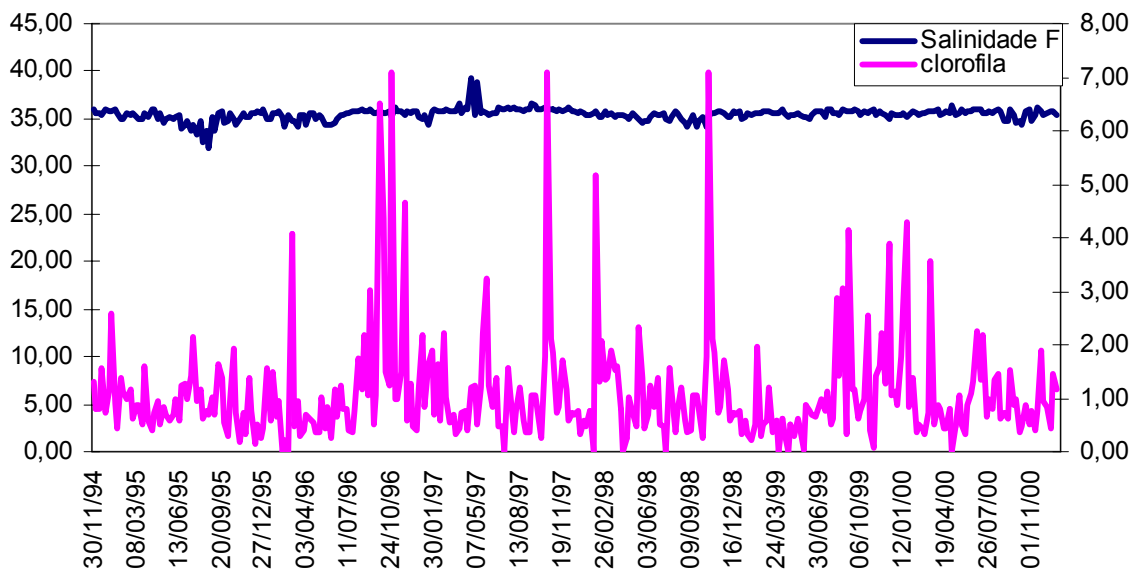


Figura II.4f – Flutuações interanuais da clorofila x salinidade no fundo.

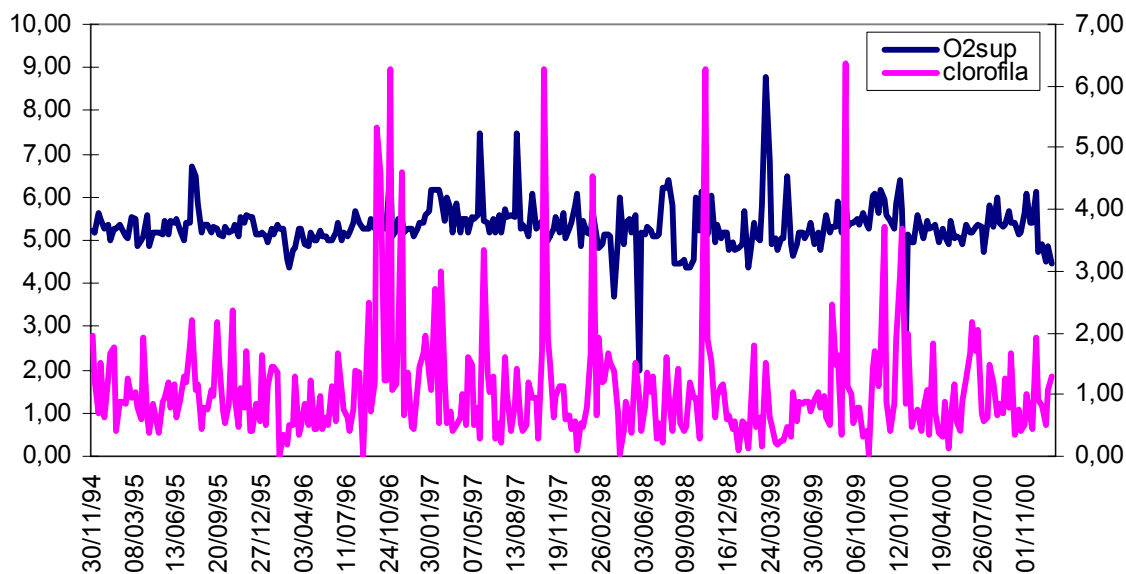


Figura II.4g – Flutuações interanuais da clorofila x oxigênio na superfície.

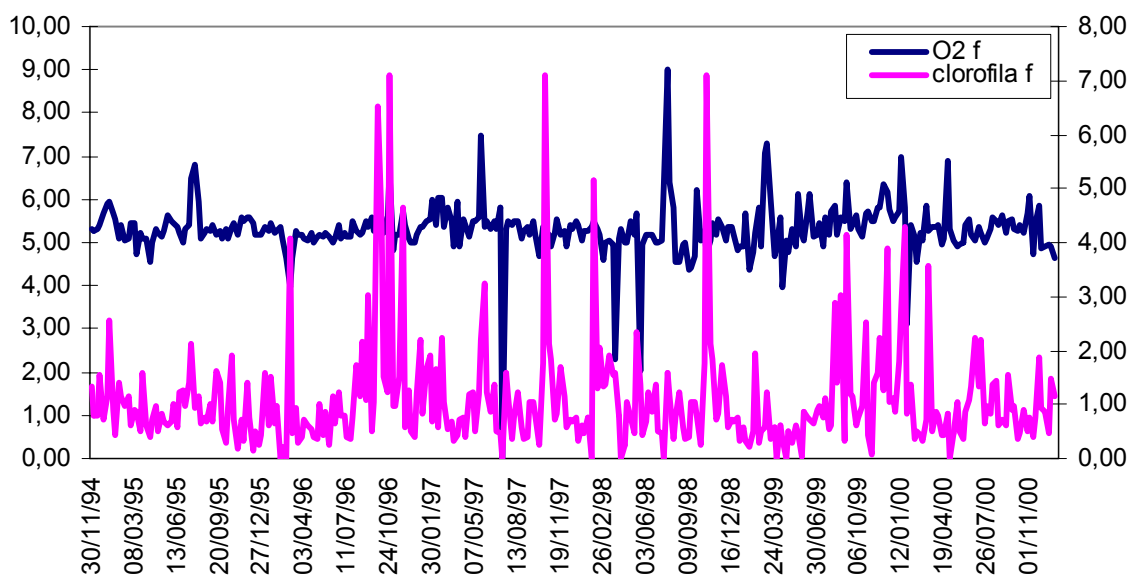


Figura II.4h – Flutuações interanuais da clorofila x oxigênio no fundo.

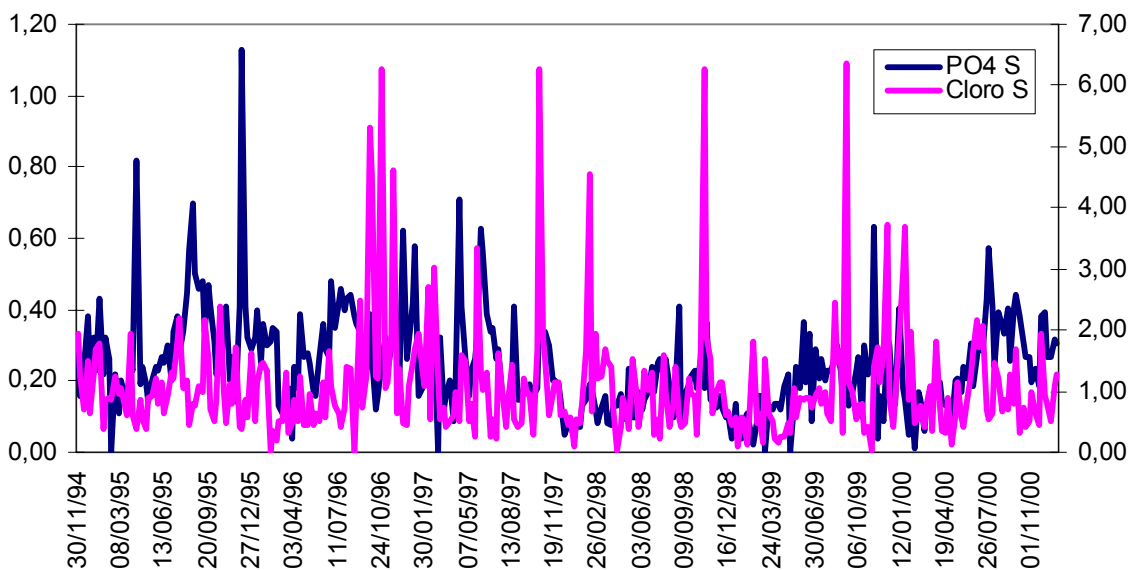


Figura II.4i – Flutuações interanuais da clorofila x fosfato inorgânico na superfície.

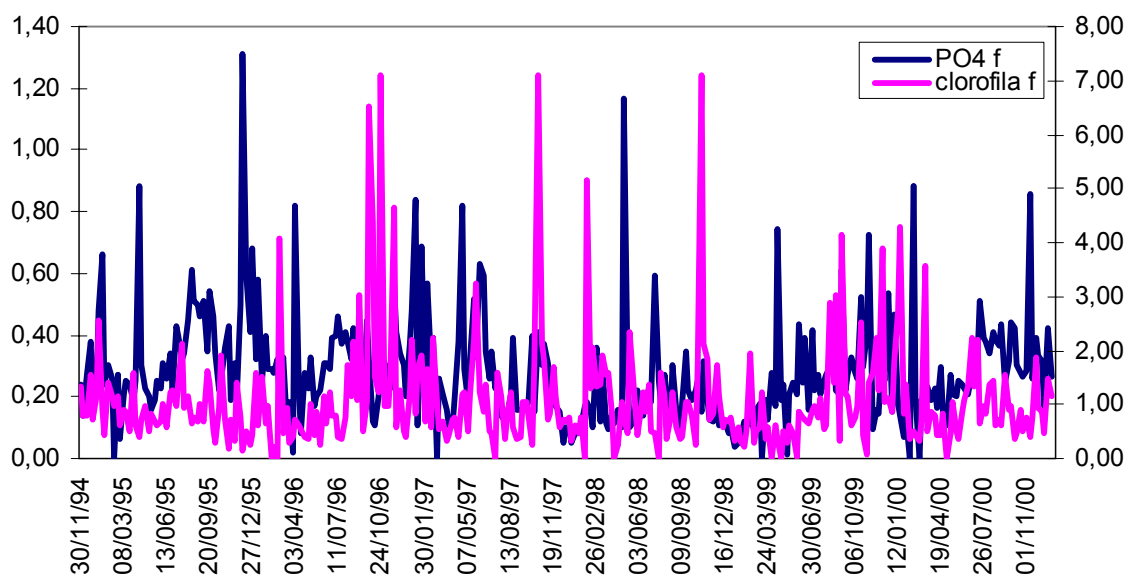


Figura II.4j – Flutuações interanuais da clorofila x fosfato inorgânico no fundo.

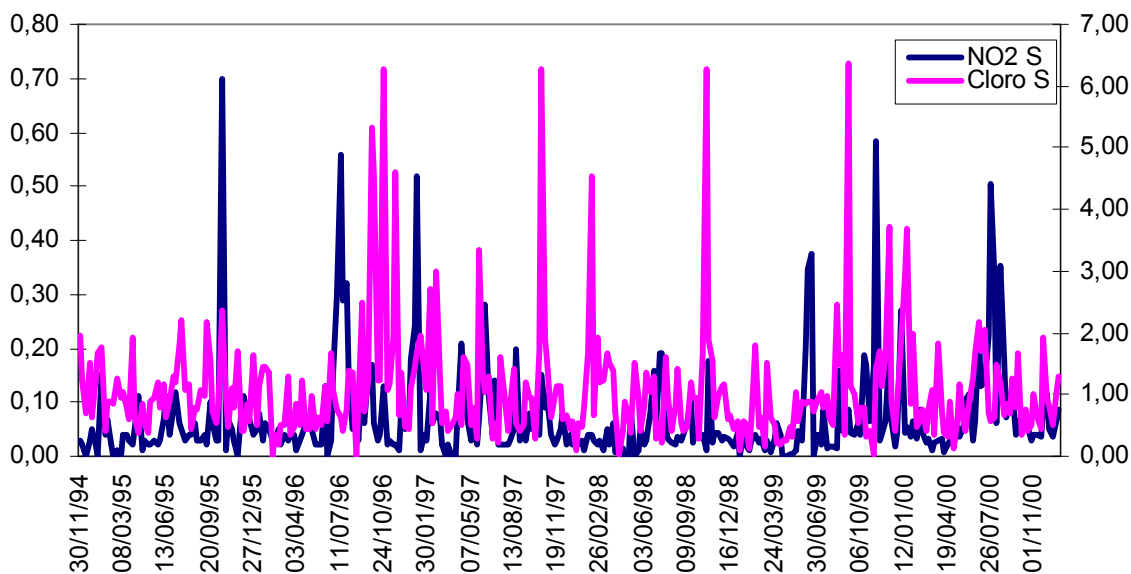


Figura II4l – Flutuações interanuais da clorofila x NO₂ na superfície.

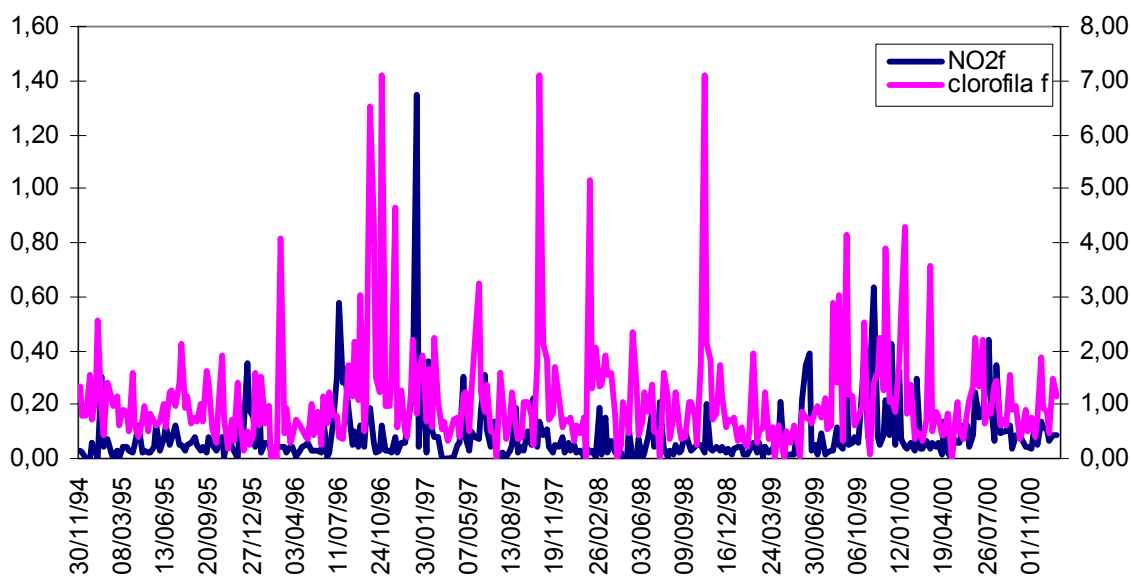


Figura II.4m – Flutuações interanuais da clorofila x NO₂ no fundo.

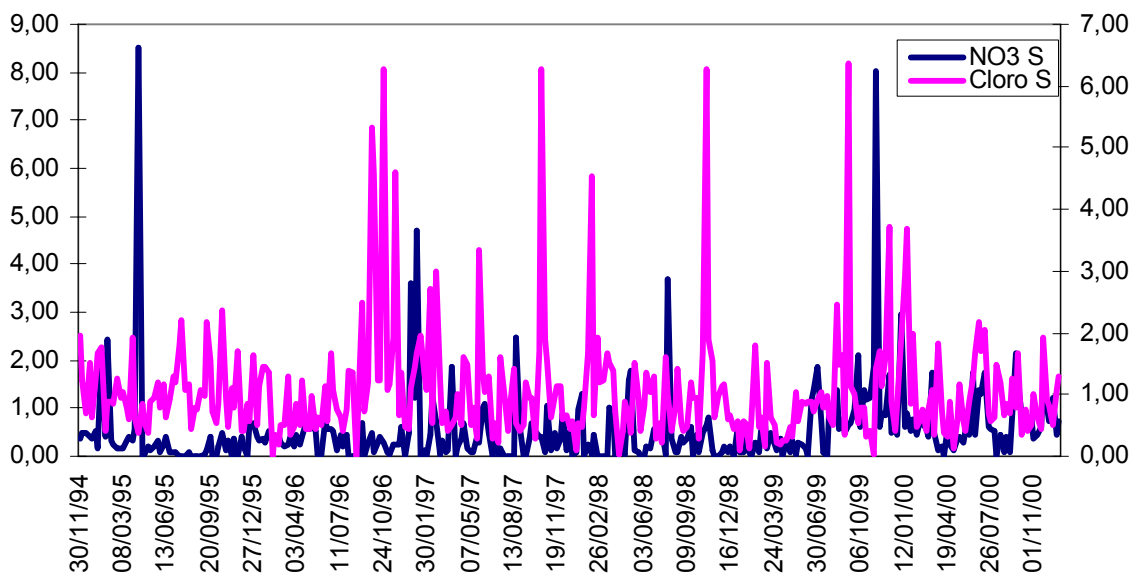


Figura II.4n – Flutuações interanuais da clorofila x NO₃ na superfície.

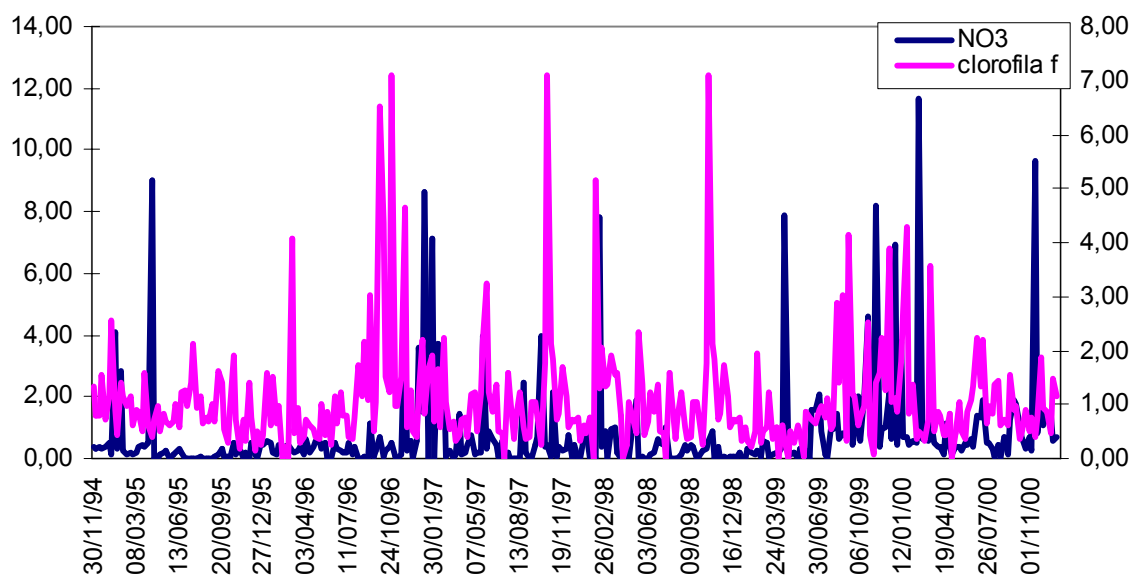


Figura II.4 o – Flutuações interanuais da clorofila x NO₃ no fundo.

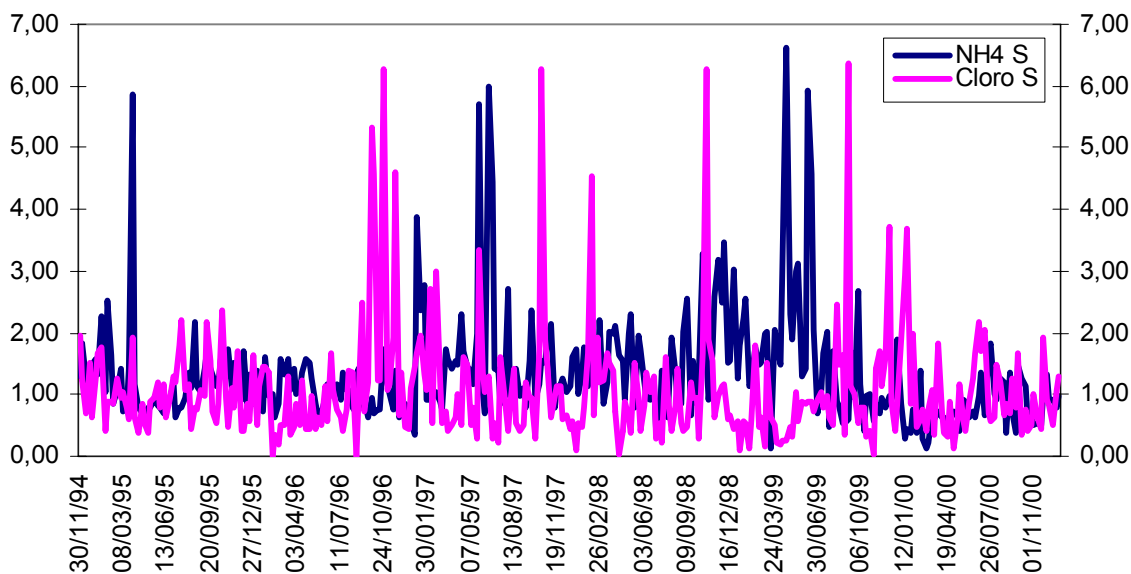


Figura II.4p – Flutuações interanuais de clorofila x NH4 na superfície

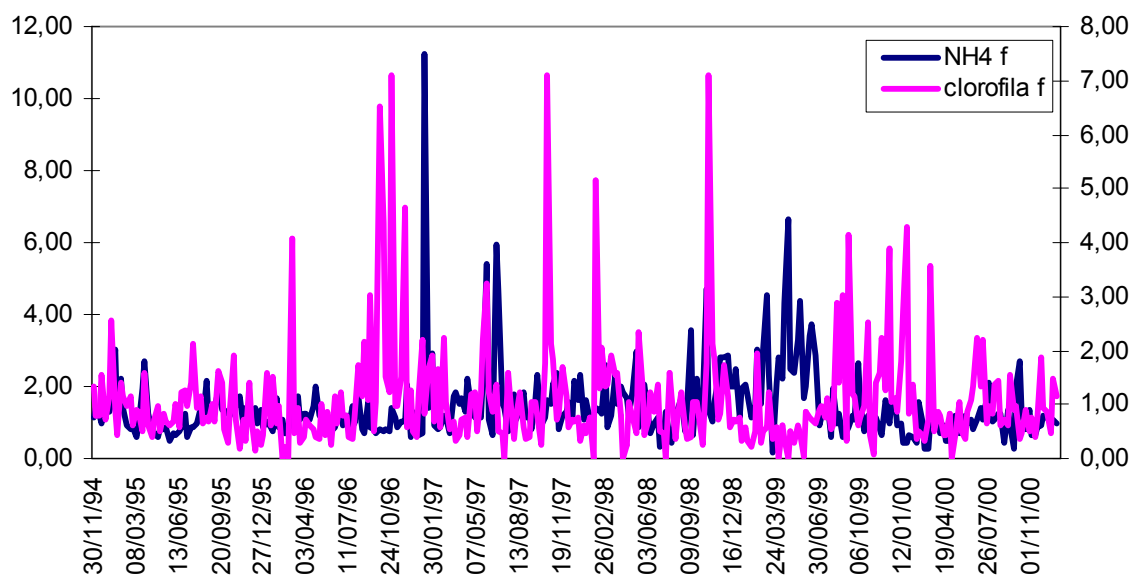


Figura II.4q – Flutuações interanuais de clorofila x NH4 no fundo.

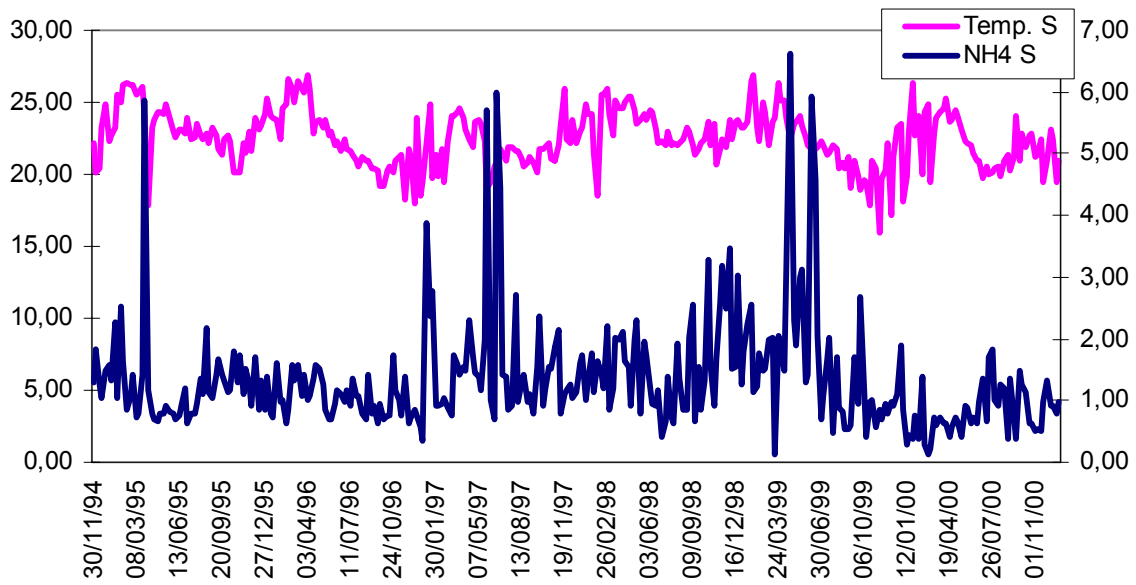


Figura II.5 a – Flutuações interanuais da temperatura x PO_4 na superfície.

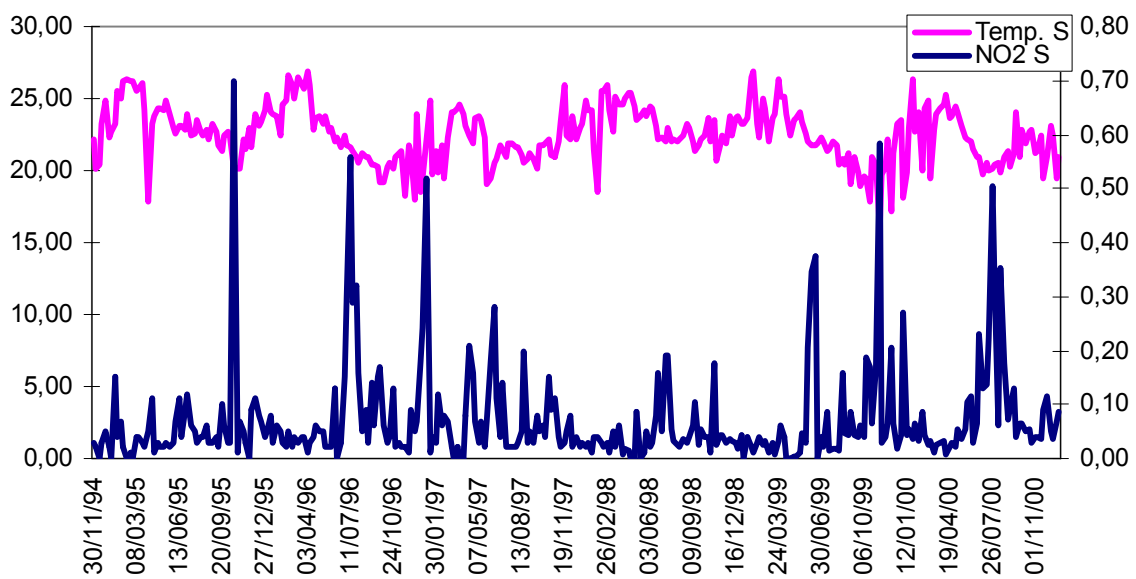


Figura II.5b – Flutuações interanuais da temperatura x NO_2 na superfície.

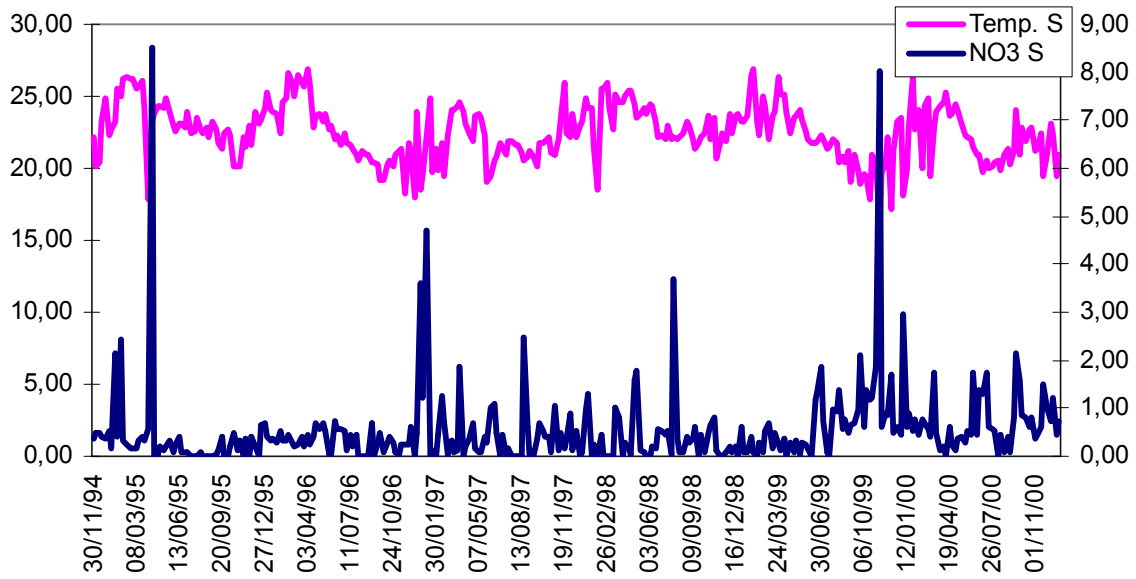


Figura II.5c – Flutuações interanuais da temperatura x NO₃ na superfície

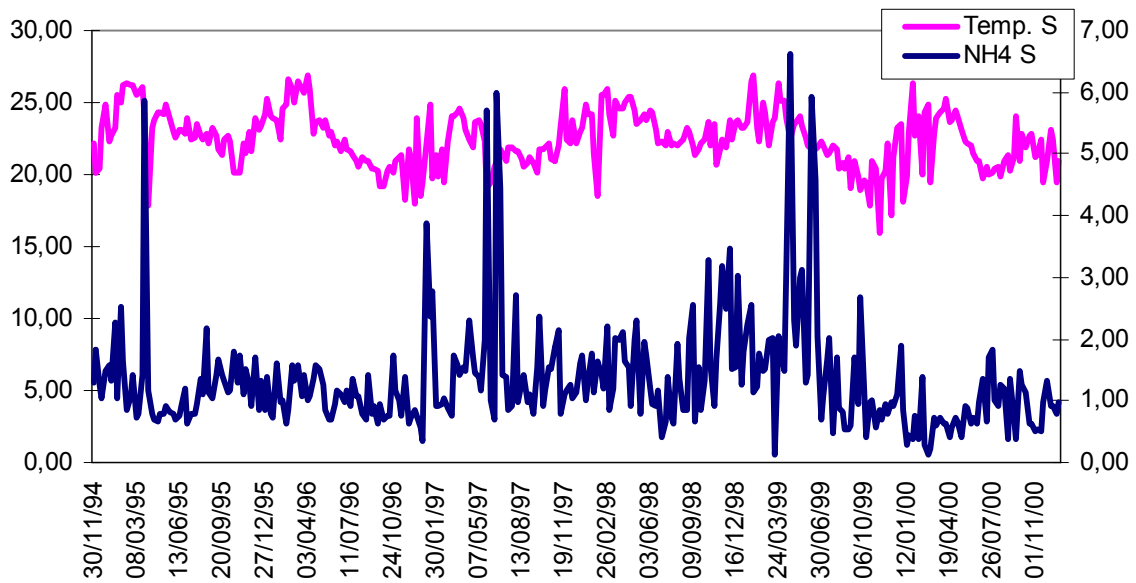


Figura 5d – Flutuações interanuais da temperatura x NH₄ na superfície.

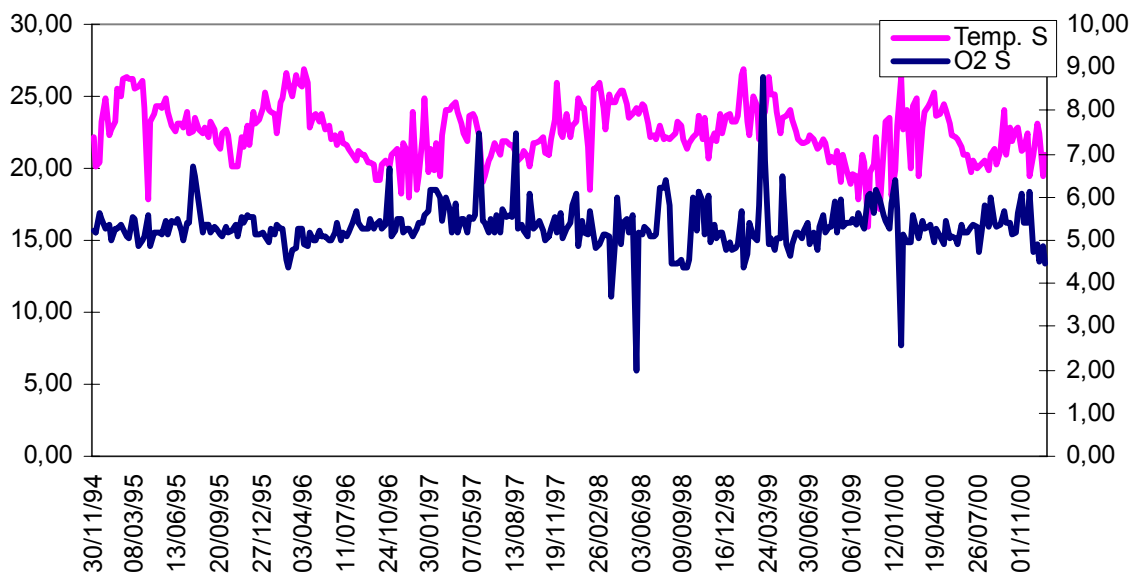


Figura II.5e – Flutuações interanuais da temperatura x O₂ na superfície.

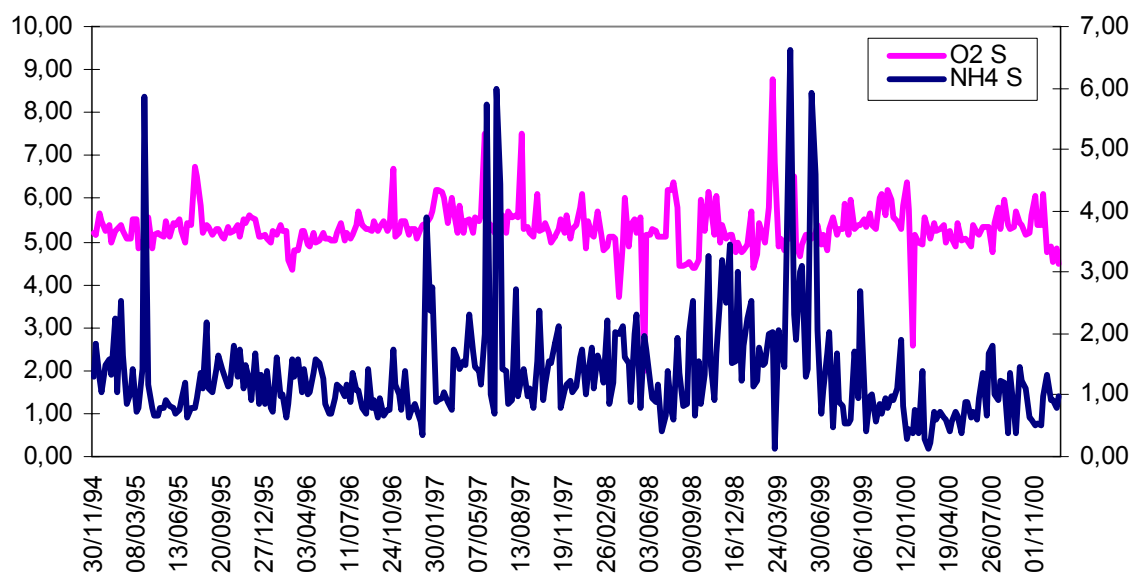


Figura II.5 f – Flutuações interanuais do O₂ x NH₄ na superfície.

Anexo III

PARÂMETROS E ARQUITETURAS DOS MODELOS DE REDES NEURAISSCOLHIDOS PARA PREVISÃO DA SÉRIE TEMPORAL DA ATIVIDADE FITOPLANCTONICA

PARÂMETROS COMUNS:

Algoritmo genético:

Máximo de gerações: 10

Tamanho da população: 50

Seleção: 50% sobrevive

Recombinação usa a técnica Tail Swap

Preenchimento da população por clonagem

Mutação: 25% por alteração aleatória

Para as redes MLP (Perceptron de Múltiplas Camadas)

Número de centros: máximo de 16

Sobreposição: de 0,4 a 1,8

Proximidade: de 1 a 3

Rede:

Ciclos mínimos para treinamento da rede: 20

Ciclos máximos para treinamento da rede: 50

Número máximo de camadas escondidas: 2

Limite máximo de neurônios na camada escondida: 16

Intervalo para inicialização dos pesos: $\pm 0,25$

Interrompe o treinamento após 15 ciclos sem melhorar o desempenho da mesma. Treina cada rede 10 vezes e calcula o resultado.

Critério de parada: 10 gerações ou RMSE menor que 0,01

Primeira camada escondida

Taxa de aprendizado: de 0,8 até 0,1

Momentum: de 0,6 até 0,1

Segunda camada escondida

Taxa de aprendizado: de 0,8 até 0,1

Momentum: de 0,6 até 0,1

Camada de Saída

Taxa de aprendizado: de 0,4 até 0,1

Momentum: de 0,2 até 0,1

ARQUITETURAS DAS REDES PARA A PREVISÃO DE CLOROFILA DE SUPERFÍCIE

MODELO 4:

Rede: Perceptron de Múltiplas Camadas (MLP)

Número de Neurônios:

Camada de entrada: 11

1ª Camada escondida: 14 (6 Sigmóide , 8 Tangente Hiperbólica)

Camada de Saída: 1(Sigmóide).

MODELO 5:

Rede: Perceptron de Múltiplas Camadas (MPL)

Número de Neurônios:

Camada de entrada: 13

1ª Camada escondida: 7 (2 Sigmóide, 5 Tangente Hiperbólica)

Camada de Saída: 1 (Sigmóide)

MODELO 6:

Rede: Perceptron de Múltiplas Camadas (MLP)

Número de Neurônios:

Camada de entrada: 14

1ª Camada escondida: 11 (6 Sigmóide, 4 Tangente Hiperbólica, 1 Linear)

Camada de Saída: 1 (Sigmóide)

ARQUITETURA DAS REDES PARA A PREVISÃO DA CLOROFILA DE FUNDO

MODELO 1:

Rede: Perceptron de Múltiplas Camadas (MLP)

Número de Neurônios:

Camada de entrada: 12

1ª Camada escondida: 4 (1 Tangente Hiperbólica, 3 Linear)

2ª Camada escondida: 5 (4Sigmóide, 1 Linear) – 1 conexão recorrente

Camada de Saída: 1 (Tangente Hiperbólica) – 1 conexão recorrente.

MODELO 2:

Rede: Perceptron de Múltiplas Camadas (MLP)

Número de Neurônios:

Camada de entrada: 10

1ª Camada escondida: 6 (3 Sigmóide, 2 Tangente Hiperbólica, 1 Linear)

Camada de Saída: 1 (Sigmóide)

MODELO 3:

Rede: Perceptron de Múltiplas Camadas (MLP)

Número de Neurônios:

Camada de entrada: 11

1ª Camada escondida: 9 (3 Sigmóide, 2 Tangente Hiperbólica, 4 Linear)

Camada de Saída: 1 (Sigmóide)

**DISTRIBUIÇÃO DAS VARIÁVEIS USADAS PELOS DIVERSOS
MODELOS**

Variáveis	Previsão de Fundo			Previsão de superfície		
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
Data	X		X	X		X
Temp S	X			X		X
Temp F	X			X	X	X
Sal S						
Sal F		X	X			
O2 S					X	X
O2 F	X			X		X
PO4 S		X				
PO4 lag S	X	X		X		
PO4 F	X		X		X	X
PO4 lag F	X					
N tot S			X	X	X	X
N tot lag S						
N tot F	X	X			X	X
N tot lag F	X				X	X
NO2 S		X			X	X
NO2 F		X	X		X	X
NO3 S			X	X		X
NO3 F	X		X	X		
NH4 S		X		X		
NH4 F			X	X	X	X
N/P S						
N/P F		X	X		X	
Cloro S	X	X	X			
Cloro lag S					X	
Cloro F				X	X	X
Cloro lag F	X	X	X	X		